(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 February ~~2002~~ (21.02.2002)

PCT

(10) International Publication Number
WO 02/14503 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/US01/41733

(22) International Filing Date: 14 August 2001 (14.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/225,152	14 August 2000 (14.08.2000)	US
60/236,428	28 September 2000 (28.09.2000)	US
60/270,520	21 February 2001 (21.02.2001)	US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and

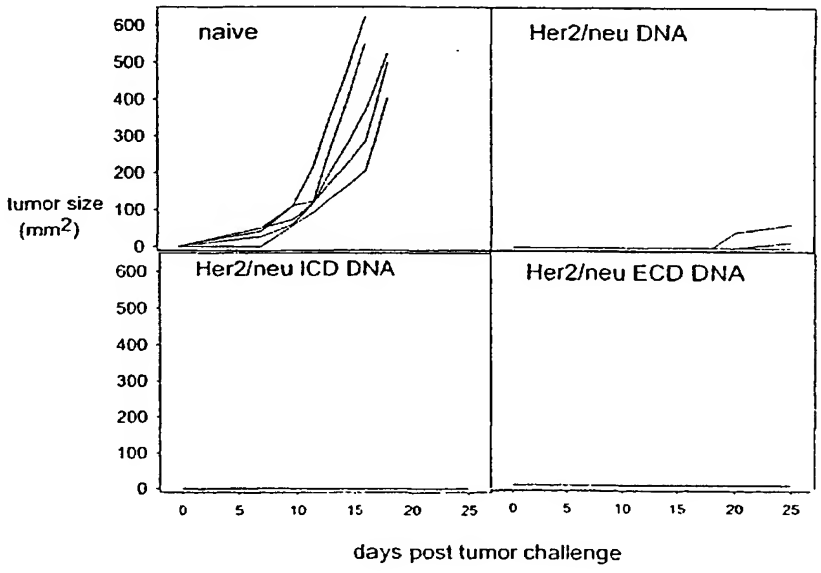
(75) Inventors/Applicants *(for US only)*: **HAND-ZIMMERMANN, Susan** [US/US]; 2014 179th Court N.E., Redmond, WA 98052 (US). **CHEEVER, Martin, A.** [US/US]; 6210 S.E. 22nd Avenue, Mercer Island, WA 98040 (US). **FOY, Teresa, M.** [US/US]; 2104 S. 277th Place, Federal Way, WA 98003 (US). **LODES, Michael, J.** [US/US]; 9223 36th Avenue S.W., Seattle, WA 98126 (US). **KALOS, Michael, D.** [US/US]; 8116 Dayton Avenue N., Seattle, WA 98103 (US). **MCNEILL, Patricia, D.** [US/US]; 1333 South 290th Place, Federal Way, WA 98003 (US). **VEDVICK, Thomas, S.** [US/US]; 124 S. 300th Place, Federal Way, WA 98003 (US).

(74) Agents: **CHRISTIANSEN, William, T.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF HER-2/NEU-ASSOCIATED MALIGNANCIES



(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly Her-2/neu-associated cancers, are disclosed. Illustrative compositions comprise one or more Her-2/neu polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of Her-2/neu-associated malignancies.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF HER-2/NEU-ASSOCIATED MALIGNANCIES

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, particularly breast cancer. The invention is more specifically related to polypeptides comprising at least an immunogenic fragment of a Her-2/Neu protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides are useful in pharmaceutical compositions, e.g., vaccines, and other compositions for the diagnosis and treatment of human malignancies.

DESCRIPTION OF THE RELATED ART

Despite enormous investments of financial and human resources, cancer remains one of the major causes of death. For example, cancer is the leading cause of death in women between the ages of 35 and 74. Breast cancer is the most common malignancy in women and the incidence for developing breast cancer is on the rise. One in nine women will be diagnosed with the disease. Standard approaches to cure breast cancer have centered around a combination of surgery, radiation and chemotherapy. These approaches have resulted in some dramatic successes in certain malignancies. However, these approaches have not been successful for all malignancies and breast cancer is most often incurable when attempting to treat beyond a certain stage. Alternative approaches to prevention and therapy are necessary.

A common characteristic of malignancies is uncontrolled cell growth. Cancer cells appear to have undergone a process of transformation from the normal phenotype to a malignant phenotype capable of autonomous growth. Amplification and overexpression of somatic cell genes is considered to be a common primary event that results in the transformation of normal cells to malignant cells. The malignant phenotypic characteristics encoded by the oncogenic genes are passed on during cell division to the progeny of the transformed cells.

Ongoing research involving oncogenes has identified at least forty oncogenes operative in malignant cells and responsible for, or associated with, transformation. Oncogenes have been classified into different groups based on the putative function or location of their gene products (such as the protein expressed by the oncogene).

Oncogenes are believed to be essential for certain aspects of normal cellular physiology. In this regard, the HER-2/*neu* oncogene is a member of the tyrosine protein kinase family of oncogenes and shares a high degree of homology with the epidermal growth factor receptor. HER-2/*neu* presumably plays a role in cell growth and/or differentiation. HER-2/*neu* appears to induce malignancies through quantitative mechanisms that result from increased or deregulated expression of an essentially normal gene product.

HER-2/*neu* (p185) is the protein product of the HER-2/*neu* oncogene. The HER-2/*neu* gene is amplified and the HER-2/*neu* protein is overexpressed in a variety of cancers including breast, ovarian, colon, lung, prostate and hematological cancers. HER-2/*neu* is related to malignant transformation. It is found in 50%-60% of ductal *in situ* carcinoma and 20%-40% of all breast cancers, as well as a substantial fraction of adenocarcinomas arising in the ovaries, prostate, colon and lung. HER-2/*neu* is intimately associated not only with the malignant phenotype, but also with the aggressiveness of the malignancy, being found in one-fourth of all invasive breast cancers. HER-2/*neu* overexpression is correlated with a poor prognosis in both breast and ovarian cancer. HER-2/*neu* is a transmembrane protein with a relative molecular mass of 185 kd that is approximately 1255 amino acids (aa) in length. It has an extracellular binding domain (ECD) of approximately 645 aa, with 40% homology to epidermal growth factor receptor (EGFR), a highly hydrophobic transmembrane anchor domain (TMD), and a carboxyterminal cytoplasmic domain (CD) of approximately 580 aa with 80% homology to EGFR.

Due to the difficulties in the current approaches to therapy of cancers in which the HER-2/*neu* oncogene is associated, there is a need in the art for improved compounds and compositions. The present invention fulfills this need, and further provides other related advantages.

BRIEF SUMMARY OF THE INVENTION

In one aspect of the present invention, Her-2/neu polypeptide and polynucleotide compositions are provided that are immunogenic, i.e., they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described herein. In one preferred embodiment, the composition is a polypeptide sequence comprising an HLA-B44 restricted, naturally processed Her-2/neu epitope, as set forth in SEQ ID NO: 3, or a polynucleotide composition encoding such a polypeptide.

The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of immunogenic activity of a polypeptide sequence set forth herein.

The present invention further provides expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, the pharmaceutical compositions, e.g., vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative

antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, pharmaceutical compositions are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions, e.g., vaccine compositions, comprising a physiologically acceptable carrier and/or an immunostimulant. The fusions proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating the expression, purification and/or immunogenicity of the polypeptide(s).

Within further aspects, the present invention provides methods for stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a Her-2/neu polynucleotide compositions, preferably a Her-2/neu polynucleotide encoding some or all of the ICD region, and more preferably a polynucleotide encoding at least the HLA-B44-restricted, naturally processed Her-2/neu epitope set forth in SEQ ID NO: 3. The patient may be afflicted with cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating $CD4^+$ and/or $CD8^+$ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer, preferably a cancer, in a patient comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the

sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Figure 1 is a graph depicting the results of ^{51}Cr-release assays demonstrating ICD reactivity in a CD8+ T cell line primed with AdV. Normal donor PBMC were primed with DC-infected with recombinant AdV expressing ICD. The assay was a standard 4 hour ^{51}Cr-release assay; targets were autologous B-LCL, either uninfected or infected with recombinant vaccinia virus expessing ICD or EGFP, as indicated. Each data point was the average of three measurements.

Figure 2 is a graph depicting the results of flow cytometric analysis of surface Her-2/neu on MCF-7 tumor cells. Cells were stained with a mAb to surface Her-2/neu, followed by a secondary rabbit anti-mouse Ig antibody conjugated to PE. Labeled cells were analyzed by flow cytometry. Values for mean fluorescent intensity were as follows: MCF-7 = 32; MCF-7 + RTV-H2N = 165; MCF-7 + Ad-H2N = 683; MCF-7 + RTV–H2N + Ad-H2N = 651.

Figure 3 illustrates that growth of EL4-Her-2/neu is inhibited by vaccination with plasmid DNA encoding Her-2/neu. Mice (5/group) were immunized (i.m.) with pVR101 Her-2/neu, pVR1012-ECD or pVR1012-ICD (100 ug) on d0 and d21. Mice were challenged

with 200, 000 EL4-Her-2/neu cells subcutaneously on d35. Tumor size was monitored for 25 days following tumor challenge.

Figure 4 illustrates that growth of EL4-Her-2/neu is partially inhibited by vaccination with the Her-2/neu ICD, but not ECD protein subunit. Mice (4/group) were immunized (s.q.) with Her-2/neu ICD or Her-2/neu ECD protein (50 ug) in Montanide 720 or d0 and d21. Mice were challenged with 200, 000 EL4-Her-2/neu cells subcutaneously on d35 Tumor size was monitored for 25 days following tumor challenge.

SEQUENCE IDENTIFIERS

SEQ ID NO: 1 sets forth a DNA sequence encoding the Her-2/neu protein.

SEQ ID NO: 2 sets forth the amino acid sequence for the Her-2/neu protein.

SEQ ID NO: 3 sets forth the amino acid sequence for a naturally processed HLA-B44-restricted epitope of Her-2/neu, corresponding to amino acids 1021-1030 of the Her-2/neu protein.

SEQ ID NO:4 is the determined cDNA for the clone HICD_CT_His_coding_region.

SEQ ID NO:5 is the determined cDNA for the clone HICD_plus_8_HIS.

SEQ ID NO:6 is the determined cDNA for the clone HICD_native_coding_region.

SEQ ID NO:7 is the determined cDNA for the clone HICD_in_pPDM_coding_sequence.

SEQ ID NO:8 is amino acid sequence encoded by the cDNA disclosed in SEQ ID NO:4.

SEQ ID NO:9 is amino acid sequence encoded by the cDNA disclosed in SEQ ID NO:6.

SEQ ID NO:10 is amino acid sequence encoded by the cDNA disclosed in SEQ ID NO:7.

SEQ ID NO:11 is amino acid sequence encoded by the cDNA disclosed in SEQ ID NO:5.

SEQ ID NO:12 is the determined cDNA for clone 68499, the TCR beta chain of the 17D5 T cell clone.

SEQ ID NO:13 is the determined cDNA for clone 68498, the TCR alpha chain of the 17D5 T cell clone.

SEQ ID NO:14 is the amino acid sequence encoded by the cDNA disclosed in SEQ ID NO:12.

SEQ ID NO:15 is the amino acid sequence encoded by the cDNA disclosed in SEQ ID NO:13.

SEQ ID NO:16 is the DNA sequence for the primer PDM-44.

SEQ ID NO:17 is the DNA sequence for the primer PDM-45.

SEQ ID NO:18 is the DNA sequence for the primer PDM-591.

SEQ ID NO:19 is the DNA sequence for the primer PDM-592.

SEQ ID NO:20 is the DNA sequence for the primer PDM-72.

SEQ ID NO:21 is the DNA sequence for the primer PDM-61.

SEQ ID NO:22 is the DNA sequence for the primer TCR Valpha-16 5'.

SEQ ID NO:23 is the DNA sequence for the primer TCR alpha 3'.

SEQ ID NO:24 is the DNA sequence for the primer TCR Vbeta-14. 5'.

SEQ ID NO:25 is the DNA sequence for the primer TCR beta 3'.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly breast cancer. As described further below, illustrative compositions of the present invention include, but are not restricted to, Her-2/neu polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (*e.g.*, T cells).

The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of virology, immunology, microbiology, molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, e.g., Sambrook, et al. Molecular

Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al. Molecular Cloning: A Laboratory Manual (1982); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Nucleic Acid Hybridization (B. Hames & S. Higgins, eds., 1985); Transcription and Translation (B. Hames & S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Perbal, A Practical Guide to Molecular Cloning (1984).

All publications, patents and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise.

HER-2/NEU POLYPEPTIDE COMPOSITIONS

As used herein, the term "polypeptide" " is used in its conventional meaning, i.e. as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, i.e. antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking an immune response.

As noted above, the present invention is directed toward compositions and methods to elicit or enhance immunity to the protein product expressed by the HER-2/*neu* oncogene, including for malignancies in a warm-blooded animal wherein an amplified HER-2/*neu* gene is associated with the malignancies. Association of an amplified HER-2/*neu* gene with a malignancy does not require that the protein

expression product of the gene be present on the tumor. For example, overexpression of the protein expression product may be involved with initiation of a tumor, but the protein expression may subsequently be lost. One embodiment of the present invention involves eliciting or enhancing an effective immune response against Her-2/neu expressing cancer cells in vivo.

More specifically, the disclosure of the present invention, in one aspect, provides polypeptides based on a particular portion (HER-2/*neu* polypeptide) of the protein expression product of the HER-2/*neu* gene can be recognized by thymus-dependent lymphocytes (hereinafter "T cells") and, therefore, an immune T cell response can be utilized prophylactically or to treat malignancies in which such a protein is or has been overexpressed.

Particularly preferred polypeptide compositions in this regard are from the ICD region of the Her-2/neu protein, preferably containing some or all of the region from about amino acids 676-1255 of SEQ ID NO: 2, and more preferably comprising at least the naturally processed HLA-B44-restricted Her-2/neu epitope set forth in SEQ ID NO: 3.

In general, CD4+ T cell populations are considered to function as helpers/inducers through the release of lymphokines when stimulated by a specific antigen; however, a subset of $CD4^+$ cells can act as cytotoxic T lymphocytes (CTL). Similarly, $CD8^+$ T cells are considered to function by directly lysing antigenic targets; however, under a variety of circumstances they can secrete lymphokines to provide helper or DTH function. Despite the potential of overlapping function, the phenotypic CD4 and CD8 markers are linked to the recognition of peptides bound to class II or class I MHC antigens. The recognition of antigen in the context of class II or class I MHC mandates that $CD4^+$ and $CD8^+$ T cells respond to different antigens or the same antigen presented under different circumstances. The binding of immunogenic peptides to class II MHC antigens most commonly occurs for antigens ingested by antigen presenting cells. Therefore, $CD4^+$ T cells generally recognize antigens that have been external to the tumor cells. By contrast, under normal circumstances, binding of peptides to class I MHC occurs only for proteins present in the cytosol and synthesized by the target itself, proteins in the external environment are excluded. An exception to this is the binding of exogenous peptides with a precise class I binding motif which are

present outside the cell in high concentration. Thus, $CD4^+$ and $CD8^+$ T cells have broadly different functions and tend to recognize different antigens as a reflection of where the antigens normally reside.

As disclosed within the present invention, a polypeptide portion of the protein product expressed by the HER-2/*neu* oncogene is recognized by T cells. Circulating HER-2/*neu* polypeptide is degraded to peptide fragments. Peptide fragments from the polypeptide bind to major histocompatibility complex (MHC) antigens. By display of a peptide bound to MHC antigen on the cell surface and recognition by host T cells of the combination of peptide plus self MHC antigen, HER-2/*neu* polypeptide (including that expressed on a malignant cell) will be immunogenic to T cells. The exquisite specificity of the T cell receptor enables individual T cells to discriminate between peptides which differ by a single amino acid residue.

During the immune response to a peptide fragment from the polypeptide, T cells expressing a T cell receptor with high affinity binding of the peptide-MHC complex will bind to the peptide-MHC complex and thereby become activated and induced to proliferate. In the first encounter with a peptide, small numbers of immune T cells will secrete lymphokines, proliferate and differentiate into effector and memory T cells. The primary immune response will occur *in vivo* but has been difficult to detect *in vitro*. Subsequent encounter with the same antigen by the memory T cell will lead to a faster and more intense immune response. The secondary response will occur either *in vivo* or *in vitro*. The *in vitro* response is easily gauged by measuring the degree of proliferation, the degree of cytokine production, or the generation of cytolytic activity of the T cell population re-exposed in the antigen. Substantial proliferation of the T cell population in response to a particular antigen is considered to be indicative of prior exposure or priming to the antigen.

Certain compounds of this invention generally comprise HER-2/*neu* polynucleotide molecules that direct the expression of such peptides, wherein the DNA molecules may be present in a viral or other delivery vector. As noted above, the polypeptides of the present invention include variants that retain the ability to stimulate an immune response. Such variants include various structural forms of the native polypeptide. Due to the presence of ionizable amino and carboxyl groups, for example,

a HER-2/*neu* polypeptide may be in the form of an acidic or basic salt, or may be in neutral form. Individual amino acid residues may also be modified by oxidation or reduction.

The present invention also includes HER-2/*neu* polypeptides with or without glycosylation. Polypeptides expressed in yeast or mammalian expression systems may be similar to or slightly different in molecular weight and glycosylation pattern than the native molecules, depending upon the expression system. For instance, expression of DNA encoding polypeptides in bacteria such as *E. coli* typically provides non-glycosylated molecules. N-glycosylation sites of eukaryotic proteins are characterized by the amino acid triplet Asn-A_1-Z, where A_1 is any amino acid except Pro, and Z is Ser or Thr. Variants of HER-2/*neu* polypeptides having inactivated N-glycosylation sites can be produced by techniques known to those of ordinary skill in the art, such as oligonucleotide synthesis and ligation or site-specific mutagenesis techniques, and are within the scope of this invention. Alternatively, N-linked glycosylation sites can be added to a HER-2/*neu* polypeptide.

A HER-2/*neu* polypeptide may generally be obtained using a genomic or cDNA clone encoding the protein. A genomic sequence that encodes full length HER-2/*neu* is shown in SEQ ID NO:1, and the deduced amino acid sequence is presented in SEQ ID NO:2. Such clones may be isolated by screening an appropriate expression library for clones that express HER-2/*neu* protein. The library preparation and screen may generally be performed using methods known to those of ordinary skill in the art, such as methods described in Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y., 1989, which is incorporated herein by reference. Briefly, a bacteriophage expression library may be plated and transferred to filters. The filters may then be incubated with a detection reagent. In the context of this invention, a "detection reagent" is any compound capable of binding to HER-2/*neu* protein, which may then be detected by any of a variety of means known to those of ordinary skill in the art. Typical detection reagents contain a "binding agent," such as Protein A, Protein G, IgG or a lectin, coupled to a reporter group. Preferred reporter groups include enzymes, substrates, cofactors, inhibitors, dyes, radionuclides, luminescent groups, fluorescent groups and biotin. More preferably, the reporter group is horseradish peroxidase, which may be

detected by incubation with a substrate such as tetramethylbenzidine or 2,2'-azino-di-3-ethylbenz-thiazoline sulfonic acid. Plaques containing genomic or cDNA sequences that express HER-2/*neu* protein are isolated and purified by techniques known to those of ordinary skill in the art. Appropriate methods may be found, for example, in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y., 1989.

In another embodiment, a polypeptide composition of the invention may also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic fragment or variant thereof.

In another embodiment of the invention, polypeptides are provided that comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide compositions set forth herein, such as those set forth in SEQ ID NOs:2-3, 8-11, and 14-15 or those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NOs:1, 4-7, and 12-13.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequences set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provide by the present invention are immunologically reactive with an antibody and/or T-cell that reacts with a polypeptide specifically set for the herein.

A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences of the invention and evaluating their immunogenic activity as described herein and/or using any of a number of techniques well known in the art.

For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other illustrative variants include variants in which a small portion (*e.g.,* 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. As described above, modifications may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics, e.g., with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, immunogenic variant or portion of a polypeptide of the invention, one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that

defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative

hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (–0.4); threonine (–0.7); serine (–0.8); tryptophan (–0.9); tyrosine (–1.3); proline (–1.6); histidine (–3.2); glutamate (–3.5); glutamine (–3.5); aspartate (–3.5); asparagine (–3.5); lysine (–3.9); and arginine (–4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (–0.4); proline (–0.5 ± 1); alanine (–0.5); histidine (–0.5); cysteine (–1.0); methionine (–1.3); valine (–1.5); leucine (–1.8); isoleucine (–1.8); tyrosine (–2.3); phenylalanine (–2.5); tryptophan (–3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or

other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

When comparing polypeptide sequences, two sequences are said to be "identical" if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Saitou, N. Nei, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these

algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment.

In one preferred approach, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known

tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene 40*:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA 83*:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S.

Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

The fusion polypeptide can comprise a polypeptide as described herein together with an unrelated immunogenic protein, such as an immunogenic protein capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute et al. *New Engl. J. Med.*, *336*:86-91, 1997).

In one preferred embodiment, the immunological fusion partner is derived from a Mycobacterium sp., such as a Mycobacterium tuberculosis-derived Ra12 fragment. Ra12 compositions and methods for their use in enhancing the expression and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; see also, Skeiky *et al.*, *Infection and Immun.* (1999) 67:3998-4007, incorporated herein by reference). C-terminal fragments of the MTB32A coding sequence express at high levels and remain as a soluble polypeptides throughout the purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous immunogenic polypeptides with which it is fused. One preferred Ra12 fusion polypeptide comprises a 14 KD C-terminal fragment corresponding to amino acid

residues 192 to 323 of MTB32A. Other preferred Ra12 polynucleotides generally comprise at least about 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide. Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium Haemophilus influenza B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemaglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the LytA gene; *Gene 43*:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to

the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology 10*:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention, when fused with this targeting signal, will associate more efficiently with MHC class II molecules and thereby provide enhanced in vivo stimulation of $CD4^+$ T-cells specific for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc. 85*:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, e.g., are at least about 90%

pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

POLYNUCLEOTIDE COMPOSITIONS

The present invention, in other aspects, provides Her-2/neu polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a polypeptide/protein of the invention or a portion thereof) or may comprise a sequence that encodes a variant or derivative, preferably and immunogenic variant or derivative, of such a sequence.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide sequence set forth in SEQ ID NO: 1, 4-7, and 12-13, complements of a polynucleotide sequence set forth in SEQ ID NO: 1, 4-7, and 12-13, and degenerate variants thereof. In certain preferred embodiments, the Her-2/neu polynucleotide sequences set forth herein encode immunogenic epitope sequences of the ICD region of the Her-2/neu protein, preferably the epitope sequence set forth in SEQ ID NO: 3.

In other related embodiments, the present invention provides polynucleotide variants having substantial identity to the sequences disclosed herein, for example those comprising at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide sequence of this invention using the methods described herein, (e.g., BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the polypeptide encoded by the variant polynucleotide is not substantially diminished relative to a polypeptide encoded by a polynucleotide sequence specifically set forth herein). The term "variants" should also be understood to encompasses homologous genes of xenogenic origin.

In additional embodiments, the present invention provides polynucleotide fragments comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17,

18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, e.g., to 60-65°C or 65-70°C.

In certain preferred embodiments, the polynucleotides described above, e.g., polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a Her-2/neu polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by

the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be "identical" if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J.*

Mol. Biol. 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical

nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the immunogenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be

obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No. 5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to "evolve" individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000

(including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in

length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered

more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION , CHARACTERIZATION AND EXPRESSION

Polynucleotides compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references).

Many template dependent processes are available to amplify a target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (e.g., *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Any of a number of other template dependent processes, many of which are variations of the PCR ™ amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent

No. 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No. WO 88/10315), including nucleic acid sequence based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl. Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) are also well-known to those of skill in the art.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.,* a tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.,* by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The

complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, amplification techniques, such as those described above, can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res. 16*:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science 269*:202-204) and automated synthesis may be

achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, any of a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem. 264*:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J. 6*:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J. 3*:1671-1680; Broglie, R. et al. (1984) *Science 224*:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ. 17*:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci. 91* :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used

to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci. 81*:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ. 20*:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation. glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction

of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell 11*:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell 22*:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. 77*:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol. 150*:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci. 85*:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol. 55*:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include, for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul. Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med. 158*:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained

intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen. San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif. 3*:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol. 12*:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc. 85*:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

ANTIBODY COMPOSITIONS, FRAGMENTS THEREOF AND OTHER BINDING AGENTS

According to another aspect, the present invention further provides binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to "specifically bind," "immunogically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or affinity of immunological binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater affinity. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and on geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of K_{off} /K_{on} enables cancellation of all parameters not related to affinity, and is thus equal to the dissociation constant K_d. See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V") regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches within the V regions of the heavy and light chains are referred to as "hypervariable regions" which are interposed between more conserved flanking stretches known as "framework regions," or "FRs". Thus the term "FR" refers to amino acid sequences

which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen, and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

Binding agents may be further capable of differentiating between patients with and without a cancer, such as breast cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of

recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol. 6*:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from

the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "$F(ab')_2$ " fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM, and on rare occasions IgG or IgA immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a non-covalent $V_H::V_L$ heterodimer including an antigen-binding site which retains much of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("sFv") polypeptide is a covalently linked $V_H::V_L$ heterodimer which is expressed from a gene fusion including V_H- and V_L-encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an sFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRS and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three

hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs, comprising the CDR set from each of a heavy and a light chain V region. A polypeptide comprising a single CDR, (e.g., a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues directly adjacent to the CDRS. Within FRs, certain amino residues and certain structural features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures--regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) Nature 349:293-299; Lobuglio et al. (1989) Proc. Nat. Acad. Sci. USA 86:4220-4224; Shaw et al. (1987) J Immunol. 138:4534-4538; and Brown et al. (1987) Cancer Res. 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) Nature 332:323-327; Verhoeyen et al. (1988) Science

239:1534-1536; and Jones et al. (1986) Nature 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody molecules which limits the duration and effectiveness of therapeutic applications of those moieties in human recipients.

As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, e.g., a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule comprising an antigen-binding site which retains substantially all of the native FR polypeptide folding structure. Veneering techniques are based on the understanding that the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the antigen-binding surface. Davies et al. (1990) Ann. Rev. Biochem. 59:439-473. Thus, antigen binding specificity can be preserved in a humanized antibody only wherein the CDR structures, their interaction with each other, and their interaction with the rest of the V region domains are carefully maintained. By using veneering techniques, exterior (e.g., solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises either a weakly immunogenic, or substantially non-immunogenic veneered surface.

The process of veneering makes use of the available sequence data for human antibody variable domains compiled by Kabat et al., in Sequences of Proteins of Immunological Interest, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V region amino acids can be deduced from the known three-dimensional structure for human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V

regions are then compared residue by residue to corresponding murine amino acids. The residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only carried out with moieties which are at least partially exposed (solvent accessible), and care is exercised in the replacement of amino acid residues which may have a significant effect on the tertiary structure of V region domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "veneered" murine antigen-binding sites are thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the residues believed to participate in non-covalent (e.g., electrostatic and hydrophobic) contacts between heavy and light chain domains, and the residues from conserved structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a murine antigen-binding site into human-appearing FRs that can be used to transfect mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y, ^{123}I, ^{125}I, ^{131}I, ^{186}Re, ^{188}Re, ^{211}At, and ^{212}Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diptheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-

containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be

coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

T CELL COMPOSITIONS

The present invention, in another aspect, provides T cells specific for a Her-2/neu polypeptide disclosed herein, or for a variant or derivative thereof. In one preferred embodiment, the T cells are specific for the Her-2/neu peptide set forth in SEQ ID NO: 3. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide.

Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor polypeptide or polynucleotide of the invention is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* *54*:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.,* by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 μg/ml, preferably 200 ng/ml - 25 μg/ml) for 3 - 7 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.,* TNF or IFN-γ) is indicative of T cell activation (*see* Coligan et al., Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a tumor polypeptide, polynucleotide or polypeptide-expressing APC may be $CD4^+$ and/or $CD8^+$. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, $CD4^+$ or $CD8^+$ T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a

short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

T CELL RECEPTOR COMPOSITIONS

The T cell receptor (TCR) consists of 2 different, highly variable polypeptide chains, termed the T-cell receptor α and β chains, that are linked by a disulfide bond (Janeway, Travers, Walport. *Immunobiology.* Fourth Ed., 148-159. Elsevier Science Ltd/Garland Publishing. 1999). The α/β heterodimer complexes with the invariant CD3 chains at the cell membrane. This complex recognizes specific antigenic peptides bound to MHC molecules. The enormous diversity of TCR specificities is generated much like immunoglobulin diversity, through somatic gene rearrangement. The β chain genes contain over 50 variable (V), 2 diversity (D), over 10 joining (J) segments, and 2 constant region segments (C). The α chain genes contain over 70 V segments, and over 60 J segments but no D segments, as well as one C segment. During T cell development in the thymus, the D to J gene rearrangement of the β chain occurs, followed by the V gene segment rearrangement to the DJ. This functional VDJ_β exon is transcribed and spliced to join to a C_β. For the α chain, a V_α gene segment rearranges to a J_α gene segment to create the functional exon that is then transcribed and spliced to the C_α. Diversity is further increased during the recombination process by the random addition of P and N-nucleotides between the V, D, and J segments of the β chain and between the V and J segments in the α chain (Janeway, Travers, Walport. *Immunobiology.* Fourth Ed., 98 and 150. Elsevier Science Ltd/Garland Publishing. 1999).

The present invention, in another aspect, provides TCRs specific for a Her-2/Neu polypeptide disclosed herein, or for a variant or derivative thereof. In particular the present invention provides the nucleic acid and amino acid sequences for the VJ or VDJ junctional sequences that determine the specificity of a given TCR. For

example, cDNA encoding a TCR specific for a Her-2/Neu peptide can be isolated from T cells specific for a Her-2/Neu polypeptide using standard molecular biological and recombinant DNA techniques.

The present invention further provides for suitable mammalian host cells, for example, non-specific T cells, that are transfected with a polynucleotide encoding TCRs specific for a Her-2/Neu polypeptide described herein, thereby rendering the host cell specific for the Her-2/Neu polypeptide. The α and β chains of the TCR may be contained on separate exression vectors or alternatively, on a single expression vector that also contains an internal ribosome entry site (IRES) for cap-independent translation of the gene downstream of the IRES. Said host cells expressing TCRs specific for the Her-2/Neu polypeptide may be used for adoptive immunotherapy of Her-2/Neu-associated malignancies as discussed in detail below.

In further aspects of the present invention, cloned TCRs specific for a Her-2/Neu polypeptide recited herein may be used in a kit for the diagnosis of Her-2/Neu-associated cancer. For example, the nucleic acid sequence or portions thereof, of Her-2/Neu-associated tumor-specific TCRs can be used as probes or primers for the detection of expression of the rearranged genes encoding for the specific TCR in a biological sample. Therefore, the present invention further provides for an assay for detecting messager RNA or DNA encoding the TCR specific for a Her-2/Neu polypeptide.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target

cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide, antibody, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and theraputic vaccine applications. Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Generally, such compositions will comprise one or more polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described herein can contain pharmaceutically acceptable salts of the polynucleotides and polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.,* salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.,* sodium, potassium, lithium, ammonium, calcium and magnesium salts).

In another embodiment, illustrative immunogenic compositions, e.g., vaccine compositions, of the present invention comprise DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems 15*:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable

promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (e.g., U.S. Pat. No. 5,219,740; Miller and Rosman (1989) BioTechniques 7:980-990; Miller, A. D. (1990) Human Gene Therapy 1:5-14; Scarpa et al. (1991) Virology 180:849-852; Burns et al. (1993) Proc. Natl. Acad. Sci. USA 90:8033-8037; and Boris-Lawrie and Temin (1993) Cur. Opin. Genet. Develop. 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) J. Virol. 57:267-274; Bett et al. (1993) J. Virol. 67:5911-5921; Mittereder et al. (1994) Human Gene Therapy 5:717-729; Seth et al. (1994) J. Virol. 68:933-940; Barr et al. (1994) Gene Therapy 1:51-58; Berkner, K. L. (1988) BioTechniques 6:616-629; and Rich et al. (1993) Human Gene Therapy 4:461-476).

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, e.g., U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebkowski et al. (1988) Molec. Cell. Biol. 8:3988-3996; Vincent et al. (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129;

Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Shelling and Smith (1994) Gene Therapy 1:165-169; and Zhou et al. (1994) J. Exp. Med. 179:1867-1875.

Additional viral vectors useful for delivering the nucleic acid molecules encoding polypeptides of the present invention by gene transfer include those derived from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The DNA encoding a polypeptide is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome. The resulting TK.sup.(-) recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, e.g., Elroy-Stein and Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747; Fuerst et al. Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox

vector is particularly desirable in human and other mammalian species since members of the Avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant Avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, e.g., WO 91/12882; WO 89/03429; and WO 92/03545.

Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. J. Biol. Chem. (1993) 268:6866-6869 and Wagner et al. Proc. Natl. Acad. Sci. USA (1992) 89:6099-6103, can also be used for gene delivery under the invention.

Additional illustrative information on these and other known viral-based delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA 86*:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci. 569*:86-103, 1989; Flexner et al., *Vaccine 8*:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques 6*:616-627, 1988; Rosenfeld et al., *Science 252*:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA 91*:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA 90*:11498-11502, 1993; Guzman et al., *Circulation 88*:2838-2848, 1993; and Guzman et al., *Cir. Res. 73*:1202-1207, 1993.

In certain embodiments, a polynucleotide may be integrated into the genome of a target cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host

cell cycle. The manner in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

In another embodiment of the invention, a polynucleotide is administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science 259*:1745-1749, 1993 and reviewed by Cohen, *Science 259*:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In still another embodiment, a composition of the present invention can be delivered via a particle bombardment approach, many of which have been described. In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such

as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN-γ, TNFα, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol. 7*:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL® adjuvants are available from Corixa Corporation (Seattle, WA; *see*, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by

Sato et al., *Science 273*:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or *Chenopodium quinoa* saponins . Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β-escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamelar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol^{R} to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL® adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL® adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 as disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Enhanzyn®) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general formula (I):

$HO(CH_2CH_2O)_n$-A-R

Wherein, n is 1-50, A is a bond or –C(O)-, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein *n* is between 1 and 50, preferably 4-24, most preferably 9; the *R* component is C_{1-50}, preferably C_4-C_{20} alkyl and most preferably C_{12} alkyl, and *A* is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-steoryl ether, polyoxyethylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12^{th} edition: entry 7717). These adjuvant molecules are described in WO 99/52549.

The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

According to another embodiment of this invention, an immunogenic composition described herein is delivered to a host via antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be

engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.,* matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature 392*:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med. 50*:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med. 4*:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNFα to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNFα, CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.,* CD54 and CD11) and costimulatory molecules (*e.g.,* CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology 75*:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.,* vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.,* a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration,

including for example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the formulation preferably provides a relatively constant level of active component release. In other embodiments, however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (*e.g.*, polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344, 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems. such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating

agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

The pharmaceutical compositions described herein may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation, is well known in the art, some of which are briefly discussed below for general purposes of illustration.

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, Nature 1997 Mar 27;386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 1998;15(3):243-84; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may

be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared is such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally. Such approaches are well known to the skilled artisan, some of which are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515

and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent 5,466,468). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents,. for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

In one embodiment, for parenteral administration in an aqueous solution, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml

of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described, *e.g.,* in U. S. Patent 5,756,353 and U. S. Patent 5,804,212. Likewise, the delivery of drugs using intranasal microparticle

resins (Takenaga *et al.*, J Controlled Release 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroetheylene support matrix is described in U. S. Patent 5,780,045.

In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively, compositions of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Lasic, Trends Biotechnol 1998 Jul;16(7):307-21; Takakura, Nippon Rinsho 1998 Mar;56(3):691-5; Chandran *et al.*, Indian J Exp Biol. 1997 Aug;35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,552,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, J Biol Chem. 1990 Sep 25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs, radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and the like, into a variety of cultured cell lines and animals. Furthermore, he use of liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

In certain embodiments, liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.,* Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

CANCER THERAPEUTIC METHODS

In further aspects of the present invention, the pharmaceutical compositions described herein may be used for the treatment of cancer, particularly for the immunotherapy of breast cancer and other Her-2/neu-associated malignancies. Within such methods, the pharmaceutical compositions described herein are administered to a patient, typically a warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as $CD8^+$ cytotoxic T lymphocytes and $CD4^+$ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow in vivo and to survive

long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever et al., *Immunological Reviews 157*:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free

survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSTIC COMPOSITIONS, METHODS AND KITS

In another embodiment, a cancer may be detected in a patient based on the presence of one or more Her-2/neu proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, the polypeptides and polynucleotides of the invention may be used as markers to indicate the presence or absence of a cancer. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a

polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μg, and preferably about 100 ng to about 1 μg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding

partner (*see, e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as breast cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In

general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1μg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a tumor protein in a biological sample. Within certain methods, a biological sample comprising $CD4^+$ and/or $CD8^+$ T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 μg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of Her-2/neu polypeptide to serve as a control. For $CD4^+$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For $CD8^+$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a Her-2/neu polypeptide in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably,

oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, *51*:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

Priming of Her-2/neu Specific CD8+ T cells using Dendritic Cells Infected with Recombinant Adenovirus

An adenovirus (AdV) vector deleted for EIA and recombinant for the intracellular domain (ICD; from about nucleotides 2026-3765 of SEQ ID NO:1) of Her-2/neu was constructed and used to infect dendritic cells (DC) obtained from a healthy donor. Priming cultures were initiated that contained AdV-ICD-infected DC as stimulators and autologous PBMC as responders. Prior to the first restimulation, the culture was enriched for CD8+ cells, and the CD8+-enriched population was restimulations with AdV-ICD infected DC. Subsequent restimulations were on autologous fibroblasts transduced with a retrovirus recombinant for the ICD. Following the fourth in vitro stimulation, the resulting T cell line was tested for ICD-specific CTL activity by a standard 4 hour ^{51}Cr-release assay. As shown in Figure 1, the bulk T cell line contained activity specific for ICD, since the line lysed autologous B-LCL infected with vaccinia-ICD, but did not lyse C-LCL infected with vaccinia-EGFP or uninfected B-LCL targets. Each data point in Figure 1 was the average of three measurements.

Following two more rounds of stimulation, the T cell line was tested for its ability to secrete γ-IFN in response to autologous fibroblasts expressing ICD. γ-IFN ELISPOT analysis was performed using the ICD-primed CD8+ T cell line as responders against autologous fibroblasts transduced with either ICD or EGFP. In this analysis, 2 x 103 fibroblasts stimulators were plated per well with 2 x 104 responding T cells per well, in triplicate. The average Elispot number for the triplicate wells were 344 on the ICD fibroblasts and 22 on the EGFP fibroblasts. Thus, the T cell line demonstrated ICD-specific γ-IFN secretion.

To investigate the class I restriction of the CD8+ ICD-specific T cell line, antibody blocking experiments were performed using antibodies specific for carious class I molecules. Stimulators were pre-incubated either with monoclonal antibody W6/32 (HLA-A, -B and –C reactive), monoclonal antibody BB123.2 (HLA-B and –C reactive) or monoclonal antibody BB7.2 (HLA-A2 specific). T cell responses were measured using a standard overnight γ-IFN Elispot assay. Responder cells were

the ICD-specific CTL line cultured in vitro for seven stimulation cycles and used at 15,000 cells per well. Stimulators were autologous fibroblasts retrovirally transduced with either ICD or EGFP and used at 2,000 cells per well. Stimulators were incubated with the indicated mAb (50μg/mL) for 20 minutes prior to being added to the assay). The assays were performed in triplicate.

As shown in Table 2, incubation of stimulator cells with either the W6/32 or BB123.2 antibodies completely blocked recognition of the ICD-transduced fibroblasts, whereas incubation with BB7.2 had no effect on γ-IFN secretion. These results indicate that the ICD-specific activity was restricted by an HLA-B or –C allele.

Table 2

HLA-class I Antibody Blocking of ICD-specific γ-IFN Secretion

Stimulators	Antibody Added			
	None	BB7.2	W6/32	BB123.2
Fibro/EGFP	3	7	3	4
Fibro/ICD	167	213	4	5

An ICD-specific clone isolated from the bulk line was expanded and further characterized for its ability to recognize full-length Her-2/neu. Additionally, monoclonal antibodies specific for HLA Class I were used to examine the HLA-restriction of the clone. The experiment was a standard, overnight γ-IFN Elispot assay. Responder cells were the ICD-specific T cell clone, 17D5. Stimulators were autologous fibroblasts either untransduced or retrovirally transduced with either EGFP, ICD or full length Her2/neu (H2N). 10,000 17D5 cells and 10,000 stimulators were used per well. Antibodies were used at 25 μg/mL in the assay. The assay was performed in triplicate, and standard deviations were between 0 and +/- 18 for triplicates.

As shown in Table 3, the clone specifically recognized autologous fibroblasts transduced with ICD or full length Her-2/neu, but not untransduced fibroblasts or fibroblasts transduced with the irrelevant antigen EGFP. Furthermore, this reactivity was completely blocked by the addition of the pan-HLA Class I

monoclonal antibody w6/32 and by a monoclonal antibody specific for HLA-B and –C alleles (BB123.2), but not by an antibody specific for HLA-A2 (BB7.2). These results indicate that this Her2/neu-specific clone was restricted by an HLA-B or –C allele, the same pattern of HLA restriction observed for the bulk cell line from which the clone was derived.

Further analyses indicated that the response was restricted by HLA-B4402. These analyses were performed by testing the ability of clone 17D5 to recognize a panel of allogeneic fibroblasts matched at different HLA-B and –C alleles and infected with AdV-ICD or AdV-EGFP. Autologous fibroblasts, either transduced with recombinant retroviruses or infected with recombinant AdV were used as controls.

Table 3

γ-IFN Elispot Assay Testing Her-2/neu Reactivity HLA-Restriction of the ICD-specific Clone 17DS

Stimulators	Blocking Antibody			
	None	W6/32	BB123.2	BB7.2
Fibros	0	0	0	0
Fibro/EGFP	0	0	1	0
Fibro/ICD	162	3	1	165
Fibros-H2N	104	0	0	98
T cells alone	0	0	0	0

The Her-2/neu specific clone was tested for its ability to recognize human tumor cells expressing Her-2/neu. The breast carcinoma cell line MCF-7 naturally expresses low levels of Her-2/neu at the cell surface and is also HLA-b4402. Upon transduction of MCF-7 with a retrovirus recombinant for Her-2/neu, surface levels of Her-2/neu increased about 5-fold as measured by flow cytometric analysis following staining with a Her-2/neu specific monoclonal antibody. Infection of MCF-7 cells with AdV-Her-2/neu resulted in a 20-fold increase of surface Her-2/neu on the tumor cells. These results are depicted in Figure 2.

The T cell clone secreted γ-IFN in response to MCF-7 cells infected with the adenovirus encoding Her-2/neu. The clone did not, however, appear to recognize MCF-7 cells or MCF-7 cells transduced with the retrovirus expressing Her-2/neu. Since the clone does recognize human fibroblasts transduced with either ICD or Her2/neu, and since the transduced fibroblasts express similar levels of protein as the transduced MCF-7 cells, it is unlikely that this result is due solely to levels of expression of the antigen.

EXAMPLE 2

Identification of an HLA-B44-Restricted, Naturally Processed Epitope of Her-2/Neu

This example describes the charaterization of the epitope recognized by one of the T cell clones described above, 17D5. This clone recognized APC expressing the ICD or full-length Her-2/neu protein. The HLA-restriction element for the clone was determined to be HLA-B4402 by using a panel of allogenic cell lines matched at various HLA alleles with the T cell clone as APC in gamma interferon Elispot assays. This was confirmed by transduction of HLA-B44-negative, Her-2/neu-positive APC with a B4402-recombinant retrovirus to confer recognition. The region of the ICD recognized by the clone was narrowed by using recombinant retroviruses expressing a series of five fragments of the ICD to transduce B44+ APC. Recognition (as demonstrated by gamma interferon release) by the clone of two of these fragments indicated that the epitope was contained within a 235 amino acid fragment beginning at position 975 in the Her-2/neu sequence. Predicted B44-binding 9mer and 10mer peptides from within this fragment were chosen and synthesized. Of the 13 peptides synthesized, one was recognized by the clone and determined to be the epitope. This has been demonstrated by gamma interferon release and TNF-alpha release assays. The sequence of this naturally processed Her-2/neu epitope is: EEYLVPQQGF (SEQ ID NO: 3), position 1021-1030 in the Her-2/neu protein sequence.

EXAMPLE 3

Her-2/neu DNA and Polypeptide Vaccination Inhibits Growth of Her-2/Neu-Expressing Tumors

Materials and Methods

Animals. 8 -12 week old female C57Bl/6 mice were obtained from Charles River Laboratories (Wilmington, MA) and maintained in our animal facility at Corixa Corporation.

Antibodies and reagents: Rat anti-murine CD4 (GK1.5) and rat anti-murine CD8 (2.43) hybridoma cell lines were obtained from ATCC. Antibody was purified from ascites fluid. Ab-5, an anti-human Her-2/neu ECD-specific antibody, was purchased from Oncogene Research Products (Cambridge, MA). Montanide 720 was purchased from Seppic Inc. (Fairfield, NJ).

Tumor cell lines: EL4, a murine thymoma originally derived from C57BL mice, was obtained from ATCC. EL4 cells were transfected with full length human Her-2/neu using a standard electroporation protocol. EL4 cells stabely expressing Her-2/neu were obtained following in vitro drug selection with neomycin. Her-2/neu expression was confirmed by flow cytometric analysis.

Her2neu vaccines: Her-2/neu plasmid DNA vaccine (pVR1012-Her-2/neu) consisted of the full length human Her-2/neu cDNA inserted into VR1012 (Vical, San Diego, CA). The ECD plasmid DNA vaccine (pVR1012-ICD) consisted of DNA encoding amino acids 1-695 of Her-2/neu and the ICD plasmid DNA vaccine (pVR1012-ECD) consisted of DNA encoding amino acids 692-1256 in VR1012. Large quantities of endotoxin free plasmid DNA were prepared using Qiagen Inc. (Valencia, CA) kit reagents and standard techniques. Plasmid DNA vaccines were delivered intramuscularly (100 ug) on d0 and d21. ICD (amino acids 676-1256) and ECD (amino acids 22-653) recombinant subunit proteins were produced at Corixa Corporation. Briefly, ECD protein was produced by stable transfection of L cells and purified using a combination of DEAE, reverse phase

HPLC, and Mono S column chromatography. ICD protein was produced in E. Coli and purified from solubilized inclusion bodies via High Q anion exchange, followed by nickel resin affinity chromatography. Recombinant protein vaccines were mixed with Montanide 720 at a 7:2 (Montanide 720:protein) ratio and delivered subcutaneously.

In vivo tumor model. To ensure a consistent source of EL4-Her-2/neu cells for tumor protection experiments, cells were expanded by in vivo passage (i.p.) and frozen in aliquots for use in individual experiments. Tumors were established using 200,000 EL4-Her-2/neu cells injected subcutaneously on the flank. Palpable tumors typically developed within 8-10 days of injection. Tumor size is expressed in mm^2 as determined by measuring the area (length x width) of the tumors with a microcaliper device.

In vivo depletion of effector T cells: Mice were immunized with plasmid DNA or protein on days 0 and 21 to generate effector T cells. CD4 and CD8 cells were depleted by i.p. administration of 100 ug/day of purified anti-CD4 or anti-CD8 antibody on d35, 38, and 42 following initiation of the experiment. Flow cytometric analysis of depleted splenocytes indicated greater than 98% depletion of the target populations.

Adoptive transfer of immune sera: Immune sera were obtained through bleeds of Her-2/neu plasmid DNA or ICD protein immunized mice. Sera from 12 individual mice from each group were pooled for transfer (i.v.) into 6 naïve recipient mice. Anti-Her-2/neu antibody titers of immune sera were assessed by ELISA prior to sera transfer.

In vitro cytokine analysis. Mice (4/group) were immunized with 100 ug pVR1012 or pVR1012-Her-2/neu (i.m) or 50 ug of ICD protein in Montanide (s.q.), or Montanide alone on d0 and d21. Two weeks following the second immunization, 2.5×10^5 spleen cells were harvested and stimulated in vitro with media alone, ICD or ECD protein (10 ug/ml). IFNγ secretion was assayed by ELISA from supernatants 48 hours following in vitro stimulation. Values represent the mean of tripicate wells for four individual mice.

Results

Her-2/neu protein subunit and plasmid DNA vaccines mediate tumor protection.

Her-2/neu vaccines consisting of either full length or truncated forms of Her-2/neu were evaluated for the ability to elicit a protective immune response against challenge with a syngeneic Her-2/neu expressing tumor cell line. C57Bl/6 mice were immunized with plasmid DNA encoding full length human Her-2/neu, ICD, or the ECD portions of Her-2/neu. Following two DNA immunizations, mice were challenged subcutaneously with EL4 cells transfected with full length human Her-2/neu (EL4-Her-2/neu) and tumor growth was monitored. In naïve mice, EL4-Her-2/neu cells formed large solid tumors within 14-20 days of subcutaneous administration. Vaccination with Her-2/neu plasmid DNA, either full length, ICD or ECD subunits, substantially inhibited the growth of the tumor cells (Figure 3). The majority of mice are completely protected from developing tumor, whereas as small portion of animals demonstrate a delay in tumor development for up to 3 weeks following tumor challenge. It is interesting to note that similar levels of tumor protection are achieved with both the truncated and the full length Her-2/neu constructs.

To determine whether protein subunit vaccines were also effective at eliciting tumor protection, mice were immunized with ICD or ECD protein plus adjuvant, challenged with EL4-Her-2/neu, and monitored for tumor growth. The results, shown in Figure 4, demonstrate that vaccination with ICD protein elicits a partially protective immune response in which both the frequency of mice developing tumor and the mean tumor size of mice bearing tumors is decreased. In this representative experiment, ICD vaccination results in complete protection of one animal, and a decrease in mean tumor size of the mice developing tumors (162 mm^2on d23). This is compared to tumor growth in 4/4 mice (mean tumor size of 527 mm^2) in the naïve group and 4/4 mice (mean tumor size of 462 mm^2) in the ECD vaccinated group.

Unexpectedly, in comparison to Her-2/neu DNA vaccination, it is clear that with the protein vaccination was not as efficacious as DNA vaccination.

In order to determine whether the protection observed in this model was Her-2/neu specific, mice were vaccinated with full length Her-2/neu, or vector control plasmid DNA, and subsequently challenged with either parental EL4 or EL4-Her-2/neu

cells. Growth of the tumors was monitored over the next 10 to 25 days. These results demonstrated that prevention of tumor growth only occurs in mice immunized with Her-2/neu plasmid DNA, suggesting that immunity to Her-2/neu is elicited and required for protection. Further evidence that tumor protection is Her-2/neu specific is provided by the observation that vaccination with Her-2/neu plasmid DNA does not prevent growth of the parental EL4 cells. Similar results were observed when ICD protein was used as the vaccine (data not shown). Taken together, these results indicate that tumor protection mediated by Her-2/neu vaccines is Her-2/neu specific.

Mechanism of tumor protection mediated by Her-2/neu protein subunits or plasmid DNA vaccines

In order to determine the nature of the immune response responsible for mediating tumor protection with Her-2/neu plasmid DNA or protein vaccination we next performed a series of in vivo depletion and adoptive transfer experiments. The first experiments were designed to evaluate the respective roles of CD4 and CD8 effector T cells. Mice were immunized twice with full length Her-2/neu or control plasmid DNA. Two weeks following the second immunization, mice were treated in vivo with anti-CD4 or anti-CD8 antibodies to deplete effector T cells. Greater than 98% CD4 or CD8 splenic T cell depletion was achieved by 3 administrations of antibody over the course of 7 days. Three days later, mice were challenged with EL4-Her-2/neu and monitored for tumor growth. Consistent with the previous experiment shown in Figure 1, complete tumor protection is observed in mice which were vaccinated with Her-2/neu plasmid DNA (untreated group). In contrast, in vivo depletion of CD4, but not CD8, effector T cells completely abrogates tumor protection mediated by Her-2/neu DNA vaccination. Similar results were obtained in adoptive transfer experiments where it was observed that adoptive transfer of CD8-, but not CD4-depleted effector T cells conferred protection against tumor challenge (data not shown). Collectively, these results suggest that protection mediated by plasmid DNA vaccination in this system is dependent upon the presence of CD4, but not CD8 effector T cells.

Similar experiments were carried out following vaccination with ICD protein to determine the roles of CD4 and CD8 T cells in the immune response elicited by this vaccine. Again, mice were immunized and boosted with ICD protein in

adjuvant, depleted of CD4 and CD8 effector T cells by in vivo antibody treatment, and subsequently challenged with EL4-Her-2/neu. The results indicated that depletion of either CD4 or CD8 T cells abrogates the partial protection obtained with ICD vaccination suggesting that both CD4 and CD8 effector T cells play a role in ICD protein-mediated tumor protection. Results of adoptive transfer experiments also indicated that both CD4 and CD8 effector cells are important in the immune response elicited by ICD protein vaccination (data not shown).

Because it is known that anti-Her-2/neu antibodies can exhibit anti-proliferative effects on tumor cells, we investigated whether antibodies elicited by either plasmid DNA or protein vaccination contributed to the observed protection. In order to address this question, mice were immunized and boosted with full length Her-2/neu DNA or ICD protein. Sera from Her-2/neu immune mice or control sera from non-immune mice were collected and then transferred into naïve mice which were then challenged with EL4-Her-2/neu. The results from Her-2/neu DNA immune sera indicated that transfer of antibody did not confer protection. These results are somewhat predictable given that the levels of anti-Her-2/neu antibodies obtained with plasmid DNA vaccination are quite low (data not shown). Similarly, transfer of anti-ICD containing sera was not protective, despite the presence of substantial titers (10,000-100,000) of anti-ICD antibody present in this sera. Taken together, these results suggest that antibody does not mediate the protection observed in this model using EL4-Her-2/neu tumor cells.

Results of these in vivo depletion and adoptive transfer experiments indicate that CD4+ T cells play a major role in the elicitation of a protective anti-tumor immune response in this model. In order to more fully elucidate the mechanism by which CD4+ T cells mediate protection, we examined the cytokine secretion profile of T cells following vaccination with either Her-2/neu DNA or ICD protein. The results, summarized in the Table below, demonstrate that upon in vitro restimulation with recombinant ICD or ECD protein, spleen cells from Her-2/neu plasmid DNA vaccinated mice secrete substantial levels of IFNγ compared to unstimulated cells. Spleen cells from ICD protein vaccinated mice also produced IFNγ in response to in vitro stimulation with ICD, but not ECD protein. The levels of IL4 and IL-5 in these same cultures were below detection, consistent with a Th1-type immune response. Taken

together, these results suggest that IFNγ may play a role in the protection mediated by Her-2/neu vaccines in this model.

Table 4

IFNγ production following Her-2/neu DNA or protein vaccination.

vaccine[a]	IFNγ (ng/ml) medium [b]	ICD	ECD
pVR1012-Her-2/neu	0.32[c]	4.17	1.27
pVR1012	0.61	0.36	0.42
ICD protein	0.31	2.16	.01
Adjuvant alone	1.30	1.23	1.35

[a]Mice (4/group) were immunized with 100 ug pVR1012 or VR1012-Her-2/neu (i.m) or 50 ug of ICD protein in Montanide (s.q.), or Montanide alone on d0 and d21.
[b]Two weeks following the second immunization, spleen cells were harvested and stimulated in vitro with media alone,ICD or ECD protein (10 ug/ml).
[c] IFNγ secretion was assayed by ELISA 48 hours following in vitro stimulation. Values represent the mean of tripicate wells for four individual mice.

EXAMPLE 4

T Cell Clone Specific for Her-2/Neu Recognizes Human Tumor Cells

A T cell clone specific for Her-2/neu was derived by priming *in vitro* with autologous dendritic cells infected with an adenovirus recombinant for the ICD of Her-2/neu, as described in Example 1. To determine the ability of this T cell clone to recognize human tumors that endogenously express Her-2/neu, the following experiments were performed.

The human tumor cell lines SKBR3 (breast carcinoma) and SKOV3(ovarian carcinoma) both overexpress Her-2/neu. The human tumor cell lines HCT-116 (colon carcinoma) and MCF-7 (breast carcinoma) express very little or no

Her-2/neu protein. Of these tumors, only MCF-7 naturally expresses HLA-B4402 (the restriction allele for this clone) as indicated by HLA typing. A retrovirus recombinant for HLA-B4402 was used to transduce SKOV3, SKBR3, and HCT-116 tumor cell lines. Flow cytometric analysis was performed to examine HLA class I, HLA-B44, and Her-2/neu expression on parental and transduced tumor cell lines and fibroblast cell line controls. Tumor cell lines or fibroblast cell lines were stained with the following FITC-labeled monoclonal antibodies: IgG (Becton Dickinson, negative control); anti-HLA class I antibody (Sigma); anti-Bw4 antibody, which binds a subgroup of HLA-B molecules, including HLA-B44 (One Lambda); anti-Her-2/neu antibody CN2, (Ab2 from Oncogene Sciences). Samples were fixed and analyzed by flow cytometry.

Results demonstrated that all cell lines tested expressed HLA-class I at the cell surface, as expected. The autologous fibroblasts and MCF-7 tumor cells, but not the parental tumor cell lines SKBR3, SKOV3, or HCT-116, expressed HLA-B44. Following transduction with HLA-B44 retrovirus, each tumor cell line gained expression of HLA-B44. As expected, SKBR3 and SKOV3 tumor cell lines both expressed Her-2/neu at the cell surface, and levels were comparable to the amount of Her-2/neu expressed by the autologous fibroblasts retrovirally transduced with Her-2/neu. In contrast, MCF-7, HCT-116, and the non-transduced fibroblasts expressed very little or no Her-2/neu at the cell surface. MCF-7 cells transduced with retrovirus recombinant for Her-2/neu expressed levels of Her-2/neu that were comparable to those expressed by SKBR3 and SKOV3.

The ability of the Her-2/neu-specific CTL clone (clone 17D5) to recognize the above cell lines was tested in IFNγ ELISAs and TNFα bioassays. Table 5 depicts the results of an IFNγ ELISA. Clone 17D5 specifically secreted IFNγ in response to autologous HLA-B4402-positive fibroblasts transduced to express Her-2/neu. Importantly, clone 17D5 specifically secreted IFNγ in response to SKBR3 and SKOV3 tumor cells transduced with HLA-B4402, but not to the HLA-B4402-negative parental, or control transduced tumor cell lines. Clone 17D5 did not recognize the HCT-116 tumor cell line transduced with HLA-B4402. This result was expected since HCT-116 cells express only very low levels of Her-2/neu. 17D5 also did not recognize the breast tumor cell line MCF-7 or MCF-7 transduced with Her-2/neu. These results are most likely explained by insufficient levels of HLA-B4402 expressed by MCF-7, since

the levels of Her-2/neu on MCF-7 cells transduced with the Her-2/neu retrovirus are similar to the levels of Her-2/neu on the corresponding transduced fibroblasts. (This could be overcome by expressing very high levels of Her-2/neu in MCF-7 cells via infection with a Her-2/neu-recombinant adenovirus.)

Table 5

IFN$_\gamma$ ELISA Demonstrating Tumor Recognition by ICD-specific T Cell Clone 17D5[1]

Stimulators	Ave O.D.[2]
FIB	0.09
H2N Fib	1.14
HCT116-EGFP	0.11
HCT116-B44	0.10
SKBR3	0.08
SKBR3-B44	1.81
SKOV3-EGFP	0.10
SKOV3-B44	1.22
MCF7	0.09
MCF7-H2N	0.09
T cells only	0.10
Media only	0.1

[1] A standard IFNγ ELISA was performed using 24 hour supernatants obtained from incubating clone 17D5 T cells with the indicated stimulators, or media alone. 17D5 T cells and stimulators were each used at 10,000 cells/well in the assay. Assay was performed in triplicate in 96-well plates. Supernatants from stimulators incubated without T cells were included as controls, with none of the values were greater than background (data not shown). Following development of the ELISA, O.D. was read at 450nm, using 570nm as a reference.

[2] Data shown are averages of O.D. readings for triplicate wells.

The results of TNFα bioassay were consistent with the results of the IFN-γ ELISA: clone 17D5 specifically secreted TNFα in response to both SKBR3 and SKOV3 when these cell lines were transduced with the HLA-B4402-expressing retroviral construct (Table 6).

Table 6

TNFα Bioassay Demonstrating Tumor Recognition by T Cell Clone 17D5[1]

T Cells + APC	Ave O.D.[2]
FIB	0.902
H2N FIB	0.327
HCT116-EGFP	1.036
HCT116-B44	0.978
SKBR3	1.057
SKBR3-B44	0.359
SKOV3-EGFP	1.070
SKOV3-B44	0.381
MCF7	1.073
MCF7-H2N	0.878
T cells only	0.995
Media only	1.038

[1]Clone 17D5 T cells (10,000 cells/well) were incubated with the indicated APC (10,000 cells /well) or in media alone in 96-well plates in triplicate. Four hour supernatants were harvested and added to the TNFα-sensitive cell line WEHI, plated at 30,000 cells/well in a 96-well plate. WEHI cells were incubated overnight with the supernatants, and alomar blue was added to 1/10 final volume per well. O.D. 570 nm – 630nm was read at 7 hours and 24 hours after addition of Alomar blue. Results shown are from the 24 hour timepoint.

[2]O.D. values are averages of triplicate wells and indicate the relative viability of the TNFα-sensitive WEHI cells, with lower values being indicative of increased cell death, therefore increased TNFα secretion.

The above results are significant because they demonstrate that CD8+ T cells primed *in vitro* using an ICD-recombinant adenovirus are capable of recognizing human tumor cells that over express Her-2/neu. Twenty to forty percent of human breast carcinomas, as well as a proportion of carcinomas of the ovary, lung, and colon over express Her-2/neu. These data support the use of ICD as a vaccine for Her-2/neu-positive tumors.

EXAMPLE 5

Expression of Human Her-2/neu HICD in *E. coli*

This example describes constructs that were made for the expression of recombinant Human Her-2/neu ICD (HICD) protein.

The open reading frame for the human ICD was PCR amplified and sub-cloned into modified pET28 vectors, for expression of recombinant protein in *E. coli*. Two constructs were made with an N-terminal histidine tag, one with a protease cleavage site, the other without. One construct was made with a C-terminal histidine tag and one with no histidine tag.

Construction of HICD_plus_8_HIS (SEQ ID NOs:5 and 11):

The ICD coding region was originally PCR amplified from the pGS10 ATG plasmid with the following primers:

PDM-44 (SEQ ID NO:16):

5'atctctggcgcgctggatgacgatgacaagaaacgacggcagcagaag

PDM-45 (SEQ ID NO:17):

5'cagggcgcgccactcgagtcattacactggcacgtccagacccag

The PCR conditions were as follows: 10µl 10X Pfu buffer (Stratagene, La Jolla, CA), 1.25µl 10mM dNTPs (Sigma, St. Louis, MO), 3µl 10µM PDM-44 oligo, 3µl 10µM PDM-45 oligo, 80µl sterile water, 2µl Pfu DNA polymerase, ~5ng pGS10ΔATG DNA. The thermocycling conditions were as follows: a single denaturation step of 96°C for 2 minutes, followed by 40 cycles of 96°C for 30 seconds, 68°C for 15 seconds, and 72°C for 6 min. 45 sec, and a final extension of 72°C for 10 minutes. Samples were kept at 4°C until further analysis. This PCR product was cloned into a modified pT7 blue plasmid which contained an eight His tag coding region in frame with a BssHII site which is included in the PDM-44 primer. The vector and PCR product were digested with BssHII and AscI. The correct construct was screened for orientation and then sequenced. This construct was then cloned into pET14b (Novagen, Madison, WI) at the NcoI and AscI sites. This construct was then cloned into a pET28b (Novagen, Madison, WI) vector at the NcoI and HindIII sites. The final construct contains an 8-histidine tag as well as an Enterokinase cleavage site.

Construction of HICD_in_pPDM_coding_sequence (SEQ ID NOs:7 and 10):

The ICD coding region was also PCR amplified from the cDNA template with the following primers:

PDM-591 (SEQ ID NO:18) 5' cacaaacgacggcagcagaagatccggaag 3'

PDM-592 (SEQ ID NO:19) 5' gcgccactcgagtcattacactggcacgtc 3'

The PCR conditions were as follows: 10µl 10X Pfu buffer (Stratagene, La Jolla, CA), 1µl 10mM dNTPs (Sigma, St. Louis, MO), 2µl each 10µM PDM-591 and -592 oligos, 83µl sterile water, 1.5µl Pfu DNA polymerase, 1 µl cDNA. The thermocycling conditions were as follows: an initial denaturation at 96°C for 2 minutes, followed by 40 cylces of 96°C for 30 seconds, 66°C for 15 seconds, and 72°C for 5 minutes. This was followed by a final extension at 72°C for 6 minutes. The PCR product was digested with XhoI and cloned into pPDM His– a modified pET28 construct which has a His tag in frame – which had been digested with Eco 72I and XhoI. The correct construct was confirmed through sequence analysis and then transformed into BLR pLys S cells for expression.

Construction of HICD_CT_His_coding_region (SEQ ID NOs:4 and 8):

The ICD coding region was PCR amplified from the cDNA template with the following primers:

PDM-72 (SEQ ID NO:20) 5' cgacttcatatgaaacgacggcagcagaagatc 3'

PDM-61 (SEQ ID NO:21)

5'ccacgtctagagaaggcgcgccatctggatcattaatgatgatgatgatgcactggcacgtccagacccaggta 3'

The PCR conditions were as follows: 10µl 10X Pfu buffer (Stratagene, La Jolla, CA), 1µl 10mM dNTPs (Sigma, St. Louis, MO), 2µl 10µM PDM-72 oligo, 2µl 10µM PDM-61 oligo, 83µl sterile water, 1.5µl Pfu DNA polymerase, 1µl cDNA. The thermocycling conditions were as follows: an initial denaturation at 96°C for 2 minutes, followed by 40 cycles of 96°C for 30 seconds, 66°C for 15 seconds, and 72°C for 5 minutes. This was followed by a final extension at 72°C for 6 minutes. The PCR product was digested with NdeI and NotI and cloned into pPDM His – a modified pET28 construct which has a His tag in frame – which had been digested with NdeI and

NotI. The correct construct was confirmed through sequence analysis and then transformed into BLR pLys S cells for expression.

Construction of HICD_native_coding_region (SEQ ID NOs:6 and 9):

The C-terminal portion of the ICD region of Human Her-2/neu was isolated from VR102 Human Her-2/neu by digesting with KpnI and AscI. This 704 bp insert was sub-cloned into the pET28HICD with the C-terminal His tag that was also digested with KpnI and AscI (this digestion removes the C-terminal His tag from the construct, which is then replaced with the 704 bp insert). The correct construct was confirmed through sequence analysis.

EXAMPLE 6

Cloning and sequencing of TCR alpha and beta chains derived from a CD8 T cell specific for her-2/neu

This example describes the cloning and sequencing of T cell receptor (TCR) alpha and beta chains from the CD8 T cell clone specific for Her-2/neu described in Example 4. Sequence analysis demonstrated that the alpha chain of the TCR belongs to the Vα16 family and the beta chain to the Vβ14. Additionally, unique diversity and joining segments (contributing to the specificity of the response) were identified.

Total mRNA from 2 x 10^6 cells from CTL clone 17D5 was isolated using Trizol reagent and cDNA was synthesized using Ready-to-go kits (Pharmacia). To determine Vα and Vβ sequences in this clone, a panel of Vα and Vβ subtype specific primers was synthesized (based on primer sequences generated by Clontech, Palo Alto, CA) and used in RT-PCR reactions with cDNA generated from each of the clones. The RT-PCR reactions demonstrated that each of the clones expressed a common Vβ sequence that corresponded to the Vβ14 subfamily. Futhermore, using cDNA generated from the clone, the Vα sequence expressed was determined to be Vα16. To clone the full TCR alpha and beta chains from clone 17D5, primers were designed that spanned the initiator and terminator-coding TCR nucleotides. The primers were as follows:

TCR Valpha-16 5'(sense) (BamHI site---Kozak--TCR alpha sequence) (SEQ ID NO:22): GGATCC---GCCGCCACC--ATGGCCTCTGCACCCATCTCGA

TCR alpha 3' (antisense) (SalI site---TCR alpha constant sequence) (SEQ ID NO:23): GTCGAC---TCAGCTGGACCACAGCCGCAG

TCR Vbeta-14. 5'(sense) (BamHI site---Kozak--TCR alpha sequence) (SEQ ID NO:24):

GGATCC---GCCGCCACC--ATGGGCCCCCAGCTCCTTGGCTA

TCR beta 3' (antisense) (SalI site---TCR beta constant sequence) (SEQ ID NO:25): GTCGAC---TCAGAAATCCTTTCTCTTGAC.

Standard 35 cycle RT-PCR reactions were established using cDNA synthesized from the CTL clone and the above primers using the proofreading thermostable polymerase PWO (Roche, Basel, Switzerland). The resultant specific bands (~850 bp for alpha and ~950 for beta) were ligated into the PCR blunt vector (Invitrogen, Carlsbad, CA) and transformed into E.coli. E.coli transformed with plasmids containing full-length alpha and beta chains were identified, and large scale preparations of the corresponding plasmids were generated. Plasmids containing full-length TCR alpha and beta chains were submitted for sequencing. The sequencing reactions demonstrated the cloning of full-length TCR alpha and beta chains. The cDNA sequences for the alpha and beta chains are disclosed in SEQ ID NOs:13 and 12, respectively, and the amino acid sequences in SEQ ID NOs: 15 and 14, respectively. BLAST searches confirmed that the Vα belongs to the Vα16 family and the Vβ to the Vβ14 family. The diversity-joining (DJ) region that contributes to the specificity of the TCR, was unique.

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is Claimed:

1. An isolated polynucleotide composition effective for eliciting an immune response in a patient, said polynucleotide encoding a polypeptide comprising an amino acid sequence consisting essentially of SEQ ID NO: 3.

2. An isolated polypeptide composition effective for eliciting an immune response, said polypeptide comprising an amino acid sequence consisting essentially of SEQ ID NO: 3.

3. A pharmaceutical composition comprising a polynucleotide according to claim 1 or a polypeptide according to claim 2, in combination with a pharmaceutically acceptable carrier.

4. The pharmaceutical composition of claim 3, further comprising an immunostimulant.

5. The pharmaceutical composition of claim 4, wherein the immunostimulant comprises an adjuvant.

6. A method for eliciting an immune response in a patient, comprising administering to a patient an effective amount of a polynucleotide according to claim 1.

7. The method according to claim 6, wherein the patient is HLA-B44 positive.

8. The method according to claim 6, wherein the patient is afflicted with breast cancer.

9. A method for eliciting an immune response in a patient, comprising administering to a patient an effective amount of a polynucleotide having a sequence from about nucleotides 2026-3765 of SEQ ID NO:1.

10. The method according to claim 9, wherein the patient is afflicted with breast cancer.

11. An isolated polynucleotide composition comprising the TCR-alpha sequence set forth in SEQ ID NO: 13.

12. An isolated polynucleotide composition comprising the TCR-beta sequence set forth in SEQ ID NO: 12.

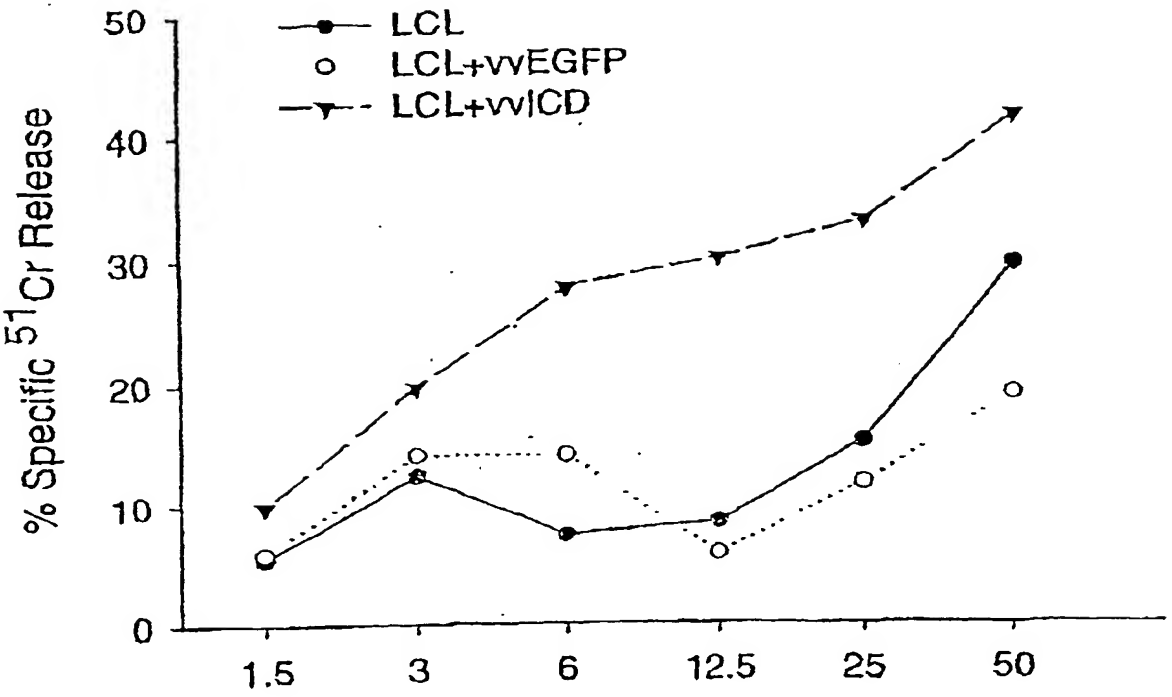

LCL
LCL+vvEGFP
LCL+vvICD
% Specific 51Cr Release
50
40
30
20
10
0
1.5
3
6
12.5
25
50

FIG. 1

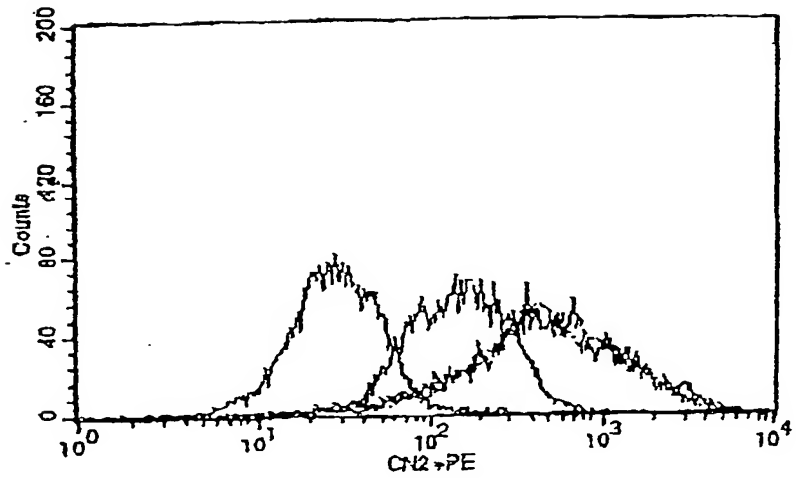


FIG. 2

FIG. 3.

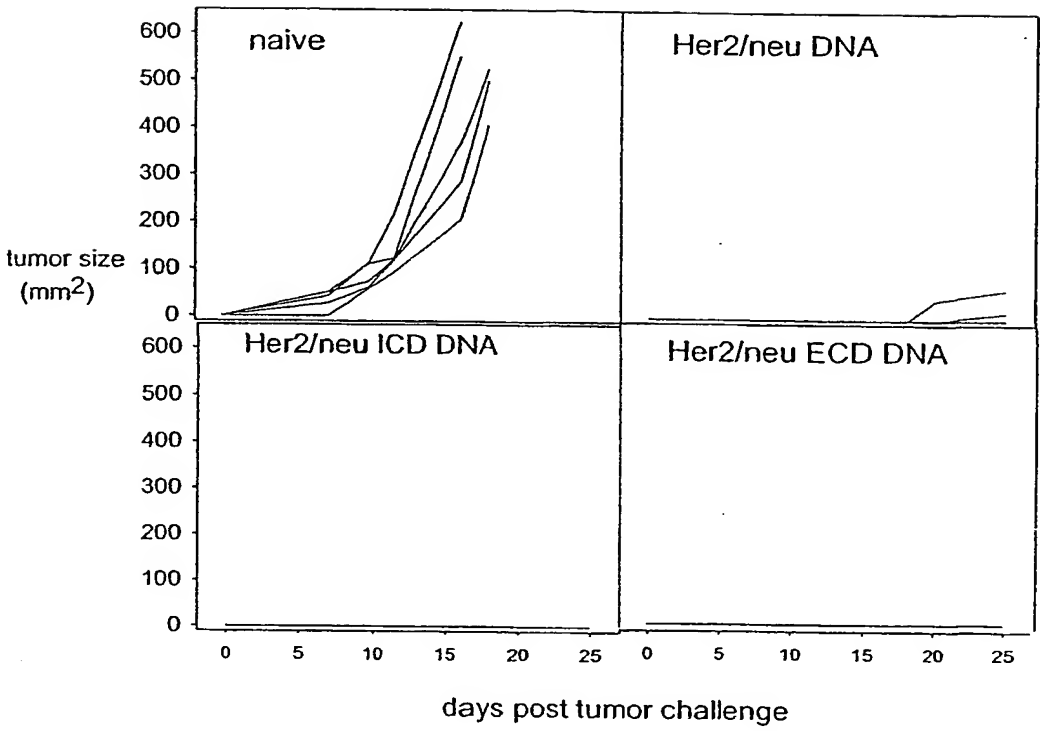
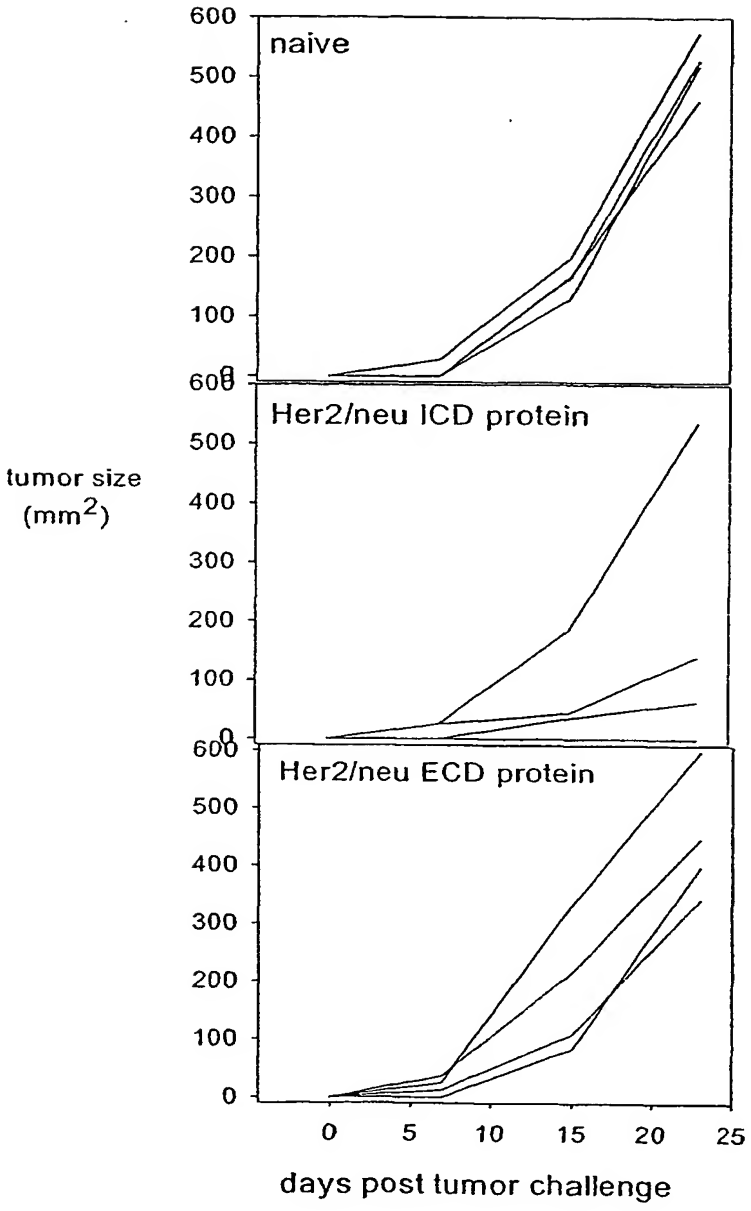

naive
Her2/neu DNA
Her2/neu ICD DNA
Her2/neu ECD DNA
tumor size (mm2)
600
500
400
300
200
100
0
0
5
10
15
20
25
days post tumor challenge

FIG. 4


naive
Her2/neu ICD protein
Her2/neu ECD protein
tumor size
(mm2)
600
500
400
300
200
100
0
0
5
10
15
20
25
days post tumor challenge

SEQUENCE LISTING

<110> Corixa Corporation
Hand-Zimmerman, Susan
Cheever, Martin A.
Foy, Teresa M.
Lodes, Michael J.
Kalos, Michael D.
McNeill, Patricia D.
Vedvick, Thomas S.

<120> COMPOSITIONS AND METHDOS FOR THE THERAPY AND DIAGNOSIS OF HER-2/NEU-ASSOCIATED MALIGNANCIES

<130> 210121.544PC

<140> PCT
<141> 2001-08-14

<160> 25

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 3768
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (1)...(3765)

<400> 1

```
atg gag ctg gcg gcc ttg tgc cgc tgg ggg ctc ctc ctc gcc ctc ttg       48
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
 1               5                  10                  15

ccc ccc gga gcc gcg agc acc caa gtg tgc acc ggc aca gac atg aag       96
Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30

ctg cgg ctc cct gcc agt ccc gag acc cac ctg gac atg ctc cgc cac      144
Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

ctc tac cag ggc tgc cag gtg gtg cag gga aac ctg gaa ctc acc tac      192
Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

ctg ccc acc aat gcc agc ctg tcc ttc ctg cag gat atc cag gag gtg      240
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

cag ggc tac gtg ctc atc gct cac aac caa gtg agg cag gtc cca ctg      288
Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

cag agg ctg cgg att gtg cga ggc acc cag ctc ttt gag gac aac tat      336
Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
```

```
                100                 105                 110

gcc ctg gcc gtg cta gac aat gga gac ccg ctg aac aat acc acc cct      384
Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

gtc aca ggg gcc tcc cca gga ggc ctg cgg gag ctg cag ctt cga agc      432
Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

ctc aca gag atc ttg aaa gga ggg gtc ttg atc cag cgg aac ccc cag      480
Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

ctc tgc tac cag gac acg att ttg tgg aag gac atc ttc cac aag aac      528
Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

aac cag ctg gct ctc aca ctg ata gac acc aac cgc tct cgg gcc tgc      576
Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

cac ccc tgt tct ccg atg tgt aag ggc tcc cgc tgc tgg gga gag agt      624
His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

tct gag gat tgt cag agc ctg acg cgc act gtc tgt gcc ggt ggc tgt      672
Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

gcc cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat gag cag tgt      720
Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

gct gcc ggc tgc acg ggc ccc aag cac tct gac tgc ctg gcc tgc ctc      768
Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

cac ttc aac cac agt ggc atc tgt gag ctg cac tgc cca gcc ctg gtc      816
His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

acc tac aac aca gac acg ttt gag tcc atg ccc aat ccc gag ggc cgg      864
Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

tat aca ttc ggc gcc agc tgt gtg act gcc tgt ccc tac aac tac ctt      912
Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

tct acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg cac aac caa      960
Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320

gag gtg aca gca gag gat gga aca cag cgg tgt gag aag tgc agc aag     1008
Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
                325                 330                 335

ccc tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac ttg cga gag     1056
Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
            340                 345                 350
```

```
gtg agg gca gtt acc agt gcc aat atc cag gag ttt gct ggc tgc aag     1104
Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
        355                 360                 365

aag atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt gat ggg gac     1152
Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
    370                 375                 380

cca gcc tcc aac act gcc ccg ctc cag cca gag cag ctc caa gtg ttt     1200
Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385                 390                 395                 400

gag act ctg gaa gag atc aca ggt tac cta tac atc tca gca tgg ccg     1248
Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
                405                 410                 415

gac agc ctg cct gac ctc agc gtc ttc cag aac ctg caa gta atc cgg     1296
Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
            420                 425                 430

gga cga att ctg cac aat ggc gcc tac tcg ctg acc ctg caa ggg ctg     1344
Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
        435                 440                 445

ggc atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg ggc agt gga     1392
Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
    450                 455                 460

ctg gcc ctc atc cac cat aac acc cac ctc tgc ttc gtg cac acg gtg     1440
Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465                 470                 475                 480

ccc tgg gac cag ctc ttt cgg aac ccg cac caa gct ctg ctc cac act     1488
Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
                485                 490                 495

gcc aac cgg cca gag gac gag tgt gtg ggc gag ggc ctg gcc tgc cac     1536
Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
            500                 505                 510

cag ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc acc cag tgt     1584
Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
        515                 520                 525

gtc aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg gag gaa tgc     1632
Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
    530                 535                 540

cga gta ctg cag ggg ctc ccc agg gag tat gtg aat gcc agg cac tgt     1680
Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545                 550                 555                 560

ttg ccg tgc cac cct gag tgt cag ccc cag aat ggc tca gtg acc tgt     1728
Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
                565                 570                 575

ttt gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac tat aag gac     1776
Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
            580                 585                 590

cct ccc ttc tgc gtg gcc cgc tgc ccc agc ggt gtg aaa cct gac ctc     1824
Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
```

```
        595                 600                 605

tcc tac atg ccc atc tgg aag ttt cca gat gag gag ggc gca tgc cag     1872
Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
    610                 615                 620

cct tgc ccc atc aac tgc acc cac tcc tgt gtg gac ctg gat gac aag     1920
Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
625                 630                 635                 640

ggc tgc ccc gcc gag cag aga gcc agc cct ctg acg tcc atc atc tct     1968
Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser
                645                 650                 655

gcg gtg gtt ggc att ctg ctg gtc gtg gtc ttg ggg gtg gtc ttt ggg     2016
Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
            660                 665                 670

atc ctc atc aag cga cgg cag cag aag atc cgg aag tac acg atg cgg     2064
Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
        675                 680                 685

aga ctg ctg cag gaa acg gag ctg gtg gag ccg ctg aca cct agc gga     2112
Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
    690                 695                 700

gcg atg ccc aac cag gcg cag atg cgg atc ctg aaa gag acg gag ctg     2160
Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
705                 710                 715                 720

agg aag gtg aag gtg ctt gga tct ggc gct ttt ggc aca gtc tac aag     2208
Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
                725                 730                 735

ggc atc tgg atc cct gat ggg gag aat gtg aaa att cca gtg gcc atc     2256
Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
            740                 745                 750

aaa gtg ttg agg gaa aac aca tcc ccc aaa gcc aac aaa gaa atc tta     2304
Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
        755                 760                 765

gac gaa gca tac gtg atg gct ggt gtg ggc tcc cca tat gtc tcc cgc     2352
Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
    770                 775                 780

ctt ctg ggc atc tgc ctg aca tcc acg gtg cag ctg gtg aca cag ctt     2400
Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
785                 790                 795                 800

atg ccc tat ggc tgc ctc tta gac cat gtc cgg gaa aac cgc gga cgc     2448
Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
                805                 810                 815

ctg ggc tcc cag gac ctg ctg aac tgg tgt atg cag att gcc aag ggg     2496
Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
            820                 825                 830

atg agc tac ctg gag gat gtg cgg ctc gta cac agg gac ttg gcc gct     2544
Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
        835                 840                 845
```

```
cgg aac gtg ctg gtc aag agt ccc aac cat gtc aaa att aca gac ttc     2592
Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
    850                 855                 860

ggg ctg gct cgg ctg ctg gac att gac gag aca gag tac cat gca gat     2640
Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865                 870                 875                 880

ggg ggc aag gtg ccc atc aag tgg atg gcg ctg gag tcc att ctc cgc     2688
Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
                885                 890                 895

cgg cgg ttc acc cac cag agt gat gtg tgg agt tat ggt gtg act gtg     2736
Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
            900                 905                 910

tgg gag ctg atg act ttt ggg gcc aaa cct tac gat ggg atc cca gcc     2784
Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
        915                 920                 925

cgg gag atc cct gac ctg ctg gaa aag ggg gag cgg ctg ccc cag ccc     2832
Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
    930                 935                 940

ccc atc tgc acc att gat gtc tac atg atc atg gtc aaa tgt tgg atg     2880
Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945                 950                 955                 960

att gac tct gaa tgt cgg cca aga ttc cgg gag ttg gtg tct gaa ttc     2928
Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
                965                 970                 975

tcc cgc atg gcc agg gac ccc cag cgc ttt gtg gtc atc cag aat gag     2976
Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
            980                 985                 990

gac ttg ggc cca gcc agt ccc ttg gac agc acc ttc tac cgc tca ctg     3024
Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
        995                 1000                1005

ctg gag gac gat gac atg ggg gac ctg gtg gat gct gag gag tat ctg     3072
Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu
    1010                1015                1020

gta ccc cag cag ggc ttc ttc tgt cca gac cct gcc ccg ggc gct ggg     3120
Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly
1025                1030                1035                1040

ggc atg gtc cac cac agg cac cgc agc tca tct acc agg agt ggc ggt     3168
Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly
                1045                1050                1055

ggg gac ctg aca cta ggg ctg gag ccc tct gaa gag gag gcc ccc agg     3216
Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg
            1060                1065                1070

tct cca ctg gca ccc tcc gaa ggg gct ggc tcc gat gta ttt gat ggt     3264
Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly
        1075                1080                1085

gac ctg gga atg ggg gca gcc aag ggg ctg caa agc ctc ccc aca cat     3312
Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His
```

```
    1090                1095                1100

gac ccc agc cct cta cag cgg tac agt gag gac ccc aca gta ccc ctg     3360
Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu
1105                1110                1115                1120

ccc tct gag act gat ggc tac gtt gcc ccc ctg acc tgc agc ccc cag     3408
Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln
                1125                1130                1135

cct gaa tat gtg aac cag cca gat gtt cgg ccc cag ccc cct tcg ccc     3456
Pro Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro
            1140                1145                1150

cga gag ggc cct ctg cct gct gcc cga cct gct ggt gcc act ctg gaa     3504
Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu
        1155                1160                1165

agg ccc aag act ctc tcc cca ggg aag aat ggg gtc gtc aaa gac gtt     3552
Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val
    1170                1175                1180

ttt gcc ttt ggg ggt gcc gtg gag aac ccc gag tac ttg aca ccc cag     3600
Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln
1185                1190                1195                1200

gga gga gct gcc cct cag ccc cac cct cct cct gcc ttc agc cca gcc     3648
Gly Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala
                1205                1210                1215

ttc gac aac ctc tat tac tgg gac cag gac cca cca gag cgg ggg gct     3696
Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala
            1220                1225                1230

cca ccc agc acc ttc aaa ggg aca cct acg gca gag aac cca gag tac     3744
Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr
        1235                1240                1245

ctg ggt ctg gac gtg cca gtg tga                                    3768
Leu Gly Leu Asp Val Pro Val
    1250                1255


      <210> 2
      <211> 1255
      <212> PRT
      <213> Homo sapien

      <400> 2
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
 1               5                  10                  15
Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30
Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45
Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80
Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95
Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
```

```
                100                 105                 110
Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125
Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140
Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160
Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175
Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190
His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205
Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220
Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240
Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255
His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270
Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285
Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300
Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320
Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
                325                 330                 335
Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
            340                 345                 350
Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
        355                 360                 365
Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
    370                 375                 380
Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385                 390                 395                 400
Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
                405                 410                 415
Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
            420                 425                 430
Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
        435                 440                 445
Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
    450                 455                 460
Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465                 470                 475                 480
Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
                485                 490                 495
Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
            500                 505                 510
Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
        515                 520                 525
Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
    530                 535                 540
Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545                 550                 555                 560
Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
                565                 570                 575
Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
            580                 585                 590
Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
```

595 600 605
Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
610 615 620
Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
625 630 635 640
Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser
645 650 655
Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
660 665 670
Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
675 680 685
Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
690 695 700
Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
705 710 715 720
Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
725 730 735
Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
740 745 750
Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
755 760 765
Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
770 775 780
Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
785 790 795 800
Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
805 810 815
Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
820 825 830
Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
835 840 845
Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
850 855 860
Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865 870 875 880
Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
885 890 895
Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
900 905 910
Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
915 920 925
Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
930 935 940
Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945 950 955 960
Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
965 970 975
Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
980 985 990
Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
995 1000 1005
Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu
1010 1015 1020
Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly
1025 1030 1035 1040
Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly
1045 1050 1055
Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg
1060 1065 1070
Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly
1075 1080 1085
Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His

```
    1090                1095                1100
Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu
1105                1110                1115                1120
Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln
                1125                1130                1135
Pro Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro
            1140                1145                1150
Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu
        1155                1160                1165
Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val
    1170                1175                1180
Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln
1185                1190                1195                1200
Gly Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala
                1205                1210                1215
Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala
            1220                1225                1230
Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr
        1235                1240                1245
Leu Gly Leu Asp Val Pro Val
    1250                1255
```

<210> 3
<211> 10
<212> PRT
<213> Homo sapiens

<400> 3
Glu Glu Tyr Leu Val Pro Gln Gln Gly Phe
1   5   10

<210> 4
<211> 1767
<212> DNA
<213> Homo sapiens

<400> 4
atgaaacgac ggcagcagaa gatccggaag tacacgatgc ggagactgct gcaggaaacg 60
gagctggtgg agccgctgac acctagcgga gcgatgccca accaggcgca gatgcggatc 120
ctgaaagaga cggagctgag gaaggtgaag gtgcttggat ctggcgcttt tggcacagtc 180
tacaagggca tctggatccc tgatggggag aatgtgaaaa ttccagtggc catcaaagtg 240
ttgagggaaa acacatcccc caaagccaac aaagaaatct tagacgaagc atacgtgatg 300
gctggtgtgg gctccccata tgtctcccgc cttctgggca tctgcctgac atccacggtg 360
cagctggtga cacagcttat gccctatggc tgcctcttag accatgtccg ggaaaaccgc 420
ggacgcctgg gctcccagga cctgctgaac tggtgtatgc agattgccaa ggggatgagc 480
tacctggagg atgtgcggct cgtacacagg gacttggccg ctcggaacgt gctggtcaag 540
agtcccaacc atgtcaaaat tacagacttc gggctggctc ggctgctgga cattgacgag 600
acagagtacc atgcagatgg gggcaaggtg cccatcaagt ggatggcgct ggagtccatt 660
ctccgccggc ggttcaccca ccagagtgat gtgtggagtt atggtgtgac tgtgtgggag 720
ctgatgactt ttggggccaa accttacgat gggatcccag cccgggagat ccctgacctg 780
ctggaaaagg gggagcggct gccccagccc cccatctgca ccattgatgt ctacatgatc 840
atggtcaaat gttggatgat tgactctgaa tgtcggccaa gattccggga gttggtgtct 900
gaattctccc gcatggccag ggacccccag cgctttgtgg tcatccagaa tgaggacttg 960
ggcccagcca gtcccttgga cagcaccttc taccgctcac tgctggagga cgatgacatg 1020
ggggacctgg tggatgctga ggagtatctg gtaccccagc agggcttctt ctgtccagac 1080
cctgccccgg gcgctggggg catggtccac cacaggcacc gcagctcatc taccaggagt 1140
ggcggtgggg acctgacact agggctggag ccctctgaag aggaggcccc caggtctcca 1200
ctggcaccct ccgaaggggc tggctccgat gtatttgatg gtgacctggg aatgggggca 1260
gccaaggggc tgcaaagcct ccccacacat gaccccagcc ctctacagcg gtacagtgag 1320
gaccccacag taccctgcc ctctgagact gatggctacg ttgccccccct gacctgcagc 1380

ccccagcctg aatatgtgaa ccagccagat gttcggcccc agccccccttc gccccgagag 1440
ggccctctgc ctgctgcccg acctgctggt gccactctgg aaaggcccaa gactctctcc 1500
ccagggaaga atggggtcgt caaagacgtt tttgcctttg ggggtgccgt ggagaacccc 1560
gagtacttga caccccaggg aggagctgcc cctcagcccc accctcctcc tgccttcagc 1620
ccagccttcg acaacctcta ttactgggac caggacccac cagagcgggg ggctccaccc 1680
agcaccttca aagggacacc tacggcagag aacccagagt acctgggtct ggacgtgcca 1740
gtgcatcatc atcatcatca ttaatga 1767

<210> 5
<211> 1806
<212> DNA
<213> Homo sapiens

<400> 5
atgggccatc atcatcatca tcatcatcat agcagcggcg cgctggatga cgatgacaag 60
aaacgacggc agcagaagat ccggaagtac acgatgcgga gactgctgca ggaaacggag 120
ctggtggagc cgctgacacc tagcggagcg atgcccaacc aggcgcagat gcggatcctg 180
aaagagacgg agctgaggaa ggtgaaggtg cttggatctg gcgcttttgg cacagtctac 240
aagggcatct ggatccctga tggggagaat gtgaaaattc cagtggccat caaagtgttg 300
agggaaaaca catcccccaa agccaacaaa gaaatcttag acgaagcata cgtgatggct 360
ggtgtgggct ccccatatgt ctcccgcctt ctgggcatct gcctgacatc cacggtgcag 420
ctggtgacac agcttatgcc ctatggctgc ctcttagacc atgtccggga aaaccgcgga 480
cgcctgggct cccaggacct gctgaactgg tgtatgcaga ttgccaaggg gatgagctac 540
ctggaggatg tgcggctcgt acacagggac ttggccgctc ggaacgtgct ggtcaagagt 600
cccaaccatg tcaaaattac agacttcggg ctggctcggc tgctggacat tgacgagaca 660
gagtaccatg cagatggggg caaggtgccc atcaagtgga tggcgctgga gtccattctc 720
cgccggcggt tcacccacca gagtgatgtg tggagttatg gtgtgactgt gtgggagctg 780
atgacttttg gggccaaacc ttacgatggg atcccagccc gggagatccc tgacctgctg 840
gaaaaggggg agcggctgcc ccagcccccc atctgcacca ttgatgtcta catgatcatg 900
gtcaaatgtt ggatgattga ctctgaatgt cggccaagat tccgggagtt ggtgtctgaa 960
ttctcccgca tggccaggga cccccagcgc tttgtggtca tccagaatga ggacttgggc 1020
ccagccagtc ccttggacag caccttctac cgctcactgc tggaggacga tgacatgggg 1080
gacctggtgg atgctgagga gtatctggta ccccagcagg gcttcttctg tccagaccct 1140
gccccgggcg ctgggggcat ggtccaccac aggcaccgca gctcatctac caggagtggc 1200
ggtggggacc tgacactagg gctggagccc tctgaagagg aggcccccag gtctccactg 1260
gcaccctccg aaggggctgg ctccgatgta tttgatggtg acctgggaat gggggcagcc 1320
aaggggctgc aaagcctccc cacacatgac cccagccctc tacagcggta cagtgaggac 1380
cccacagtac ccctgccctc tgagactgat ggctacgttg ccccccgac ctgcagcccc 1440
cagcctgaat atgtgaacca gccagatgtt cggccccagc cccttcgcc ccgagagggc 1500
cctctgcctg ctgcccgacc tgctggtgcc actctggaaa gggccaagac tctctcccca 1560
gggaagaatg gggtcgtcaa agacgttttt gcctttgggg gtgccgtgga gaaccccgag 1620
tacttgacac cccagggagg agctgcccct cagccccacc ctcctcctgc cttcagccca 1680
gccttcgaca acctctatta ctgggaccag gacccaccag agcgggggc tccacccagc 1740
accttcaaag ggacacctac ggcagagaac ccagagtacc tgggtctgga cgtgccagtg 1800
taatga 1806

<210> 6
<211> 1755
<212> DNA
<213> Homo sapiens

<400> 6
atgaaacgac ggcagcagaa gatccggaag tacacgatgc ggagactgct gcaggaaacg 60
gagctggtgg agccgctgac acctagcgga gcgatgccca accaggcgca gatgcggatc 120
ctgaaagaga cggagctgag gaaggtgaag gtgcttggat ctggcgcttt tggcacagtc 180
tacaagggca tctggatccc tgatggggag aatgtgaaaa ttccagtggc catcaaagtg 240
ttgagggaaa acacatcccc caaagccaac aaagaaatct tagacgaagc atacgtgatg 300
gctggtgtgg gctccccata tgtctcccgc cttctgggca tctgcctgac atccacggtg 360
cagctggtga cacagcttat gccctatggc tgcctcttag accatgtccg ggaaaaccgc 420
ggacgcctgg gctcccagga cctgctgaac tggtgtatgc agattgccaa ggggatgagc 480
tacctggagg atgtgcggct cgtacacagg gacttggccg ctcggaacgt gctggtcaag 540
agtcccaacc atgtcaaaat tacagacttc gggctggctc ggctgctgga cattgacgag 600

acagagtacc atgcagatgg gggcaaggtg cccatcaagt ggatggcgct ggagtccatt 660
ctccgccggc ggttcaccca ccagagtgat gtgtggagtt atggtgtgac tgtgtgggag 720
ctgatgactt ttggggccaa accttacgat gggatcccag cccgggagat ccctgacctg 780
ctggaaaagg gggagcggct gccccagccc cccatctgca ccattgatgt ctacatgatc 840
atggtcaaat gttggatgat tgactctgaa tgtcggccaa gattccggga gttggtgtct 900
gaattctccc gcatggccag ggacccccag cgctttgtgg tcatccagaa tgaggacttg 960
ggcccagcca gtcccttgga cagcaccttc taccgctcac tgctggagga cgatgacatg 1020
ggggacctgg tggatgctga ggagtatctg gtaccccagc agggcttctt ctgtccagac 1080
cctgccccgg gcgctggggg catggtccac cacaggcacc gcagctcatc taccaggagt 1140
ggcggtgggg acctgacact agggctggag ccctctgaag aggaggcccc caggtctcca 1200
ctggcaccct ccgaaggggc tggctccgat gtatttgatg gtgacctggg aatgggggca 1260
gccaaggggc tgcaaagcct ccccacacat gaccccagcc ctctacagcg gtacagtgag 1320
gaccccacag taccccctgcc ctctgagact gatggctacg ttgcccccct gacctgcagc 1380
ccccagcctg aatatgtgaa ccagccagat gttcggcccc agccccccttc gccccgagag 1440
ggccctctgc ctgctgcccg acctgctggt gccactctgg aaaggcccaa gactctctcc 1500
ccagggaaga atggggtcgt caaagacgtt tttgcctttg ggggtgccgt ggagaacccc 1560
gagtacttga caccccaggg aggagctgcc cctcagcccc accctcctcc tgccttcagc 1620
ccagccttcg acaacctcta ttactgggac caggacccac cagagcgggg ggctccaccc 1680
agcaccttca aagggacacc tacggcagag aacccagagt acctgggtct ggacgtgcca 1740
gtgtaatgac tcgag 1755

<210> 7
<211> 1773
<212> DNA
<213> Homo sapiens

<400> 7
atgcagcatc accaccatca ccaccacaaa cgacggcagc agaagatccg gaagtacacg 60
atgcggagac tgctgcagga aacggagctg gtggagccgc tgacacctag cggagcgatg 120
cccaaccagg cgcagatgcg gatcctgaaa gagacggagc tgaggaaggt gaaggtgctt 180
ggatctggcg cttttggcac agtctacaag ggcatctgga tccctgatgg ggagaatgtg 240
aaaattccag tggccatcaa agtgttgagg gaaaacacat cccccaaagc caacaaagaa 300
atcttagacg aagcatacgt gatggctggt gtgggctccc catatgtctc ccgccttctg 360
ggcatctgcc tgacatccac ggtgcagctg gtgacacagc ttatgcccta tggctgcctc 420
ttagaccatg tccgggaaaa ccgcggacgc ctgggctccc aggacctgct gaactggtgt 480
atgcagattg ccaaggggat gagctacctg gaggatgtgc ggctcgtaca cagggacttg 540
gccgctcgga acgtgctggt caagagtccc aaccatgtca aaattacaga cttcgggctg 600
gctcggctgc tggacattga cgagacagag taccatgcag atgggggcaa ggtgcccatc 660
aagtggatgg cgctggagtc cattctccgc cggcggttca cccaccagag tgatgtgtgg 720
agttatggtg tgactgtgtg ggagctgatg acttttgggg ccaaacctta cgatgggatc 780
ccagcccggg agatccctga cctgctggaa aaggggagc ggctgcccca gccccccatc 840
tgcaccattg atgtctacat gatcatggtc aaatgttgga tgattgactc tgaatgtcgg 900
ccaagattcc gggagttggt gtctgaattc tcccgcatgg ccagggaccc ccagcgcttt 960
gtggtcatcc agaatgagga cttgggccca gccagtccct tggacagcac cttctaccgc 1020
tcactgctgg aggacgatga catgggggac ctggtggatg ctgaggagta tctggtaccc 1080
cagcagggct tcttctgtcc agaccctgcc ccgggcgctg ggggcatggt ccaccacagg 1140
caccgcagct catctaccag gagtggcggt ggggacctga cactagggct ggagccctct 1200
gaagaggagg cccccaggtc tccactggca ccctccgaag gggctggctc cgatgtattt 1260
gatggtgacc tgggaatggg ggcagccaag gggctgcaaa gcctccccac acatgacccc 1320
agccctctac agcggtacag tgaggacccc acagtacccc tgccctctga gactgatggc 1380
tacgttgccc ccctgacctg cagcccccag cctgaatatg tgaaccagcc agatgttcgg 1440
ccccagcccc cttcgccccg agagggccct ctgcctgctg cccgacctgc tggtgccact 1500
ctggaaaggc ccaagactct ctccccaggg aagaatgggg tcgtcaaaga cgtttttgcc 1560
tttgggggtg ccgtggagaa ccccgagtac ttgacacccc agggaggagc tgcccctcag 1620
ccccaccctc ctcctgcctt cagcccagcc ttcgacaacc tctattactg ggaccaggac 1680
ccaccagagc gggggctccc acccagcacc ttcaaaggga cacctacggc agagaaccca 1740
gagtacctgg gtctggacgt gccagtgtaa tga 1773

<210> 8
<211> 587
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg Arg Leu
                5                  10                  15
Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Ala Met
            20                  25                  30
Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu Arg Lys
        35                  40                  45
Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Ile
    50                  55                  60
Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile Lys Val
65                  70                  75                  80
Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu
                85                  90                  95
Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg Leu Leu
            100                 105                 110
Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu Met Pro
        115                 120                 125
Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg Leu Gly
    130                 135                 140
Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly Met Ser
145                 150                 155                 160
Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn
                165                 170                 175
Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly Leu
            180                 185                 190
Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp Gly Gly
        195                 200                 205
Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg Arg Arg
    210                 215                 220
Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu
225                 230                 235                 240
Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala Arg Glu
                245                 250                 255
Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile
            260                 265                 270
Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile Asp
        275                 280                 285
Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe Ser Arg
    290                 295                 300
Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu Asp Leu
305                 310                 315                 320
Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu Leu Glu
                325                 330                 335
Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu Val Pro
            340                 345                 350
Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly Gly Met
        355                 360                 365
Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly Gly Asp
    370                 375                 380
Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg Ser Pro
385                 390                 395                 400
Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly Asp Leu
                405                 410                 415
Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His Asp Pro
            420                 425                 430
Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu Pro Ser
        435                 440                 445
Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln Pro Glu
    450                 455                 460
Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro Arg Glu
465                 470                 475                 480
```

```
Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Pro
                485                 490                 495
Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala
            500                 505                 510
Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly
        515                 520                 525
Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp
    530                 535                 540
Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro Pro
545                 550                 555                 560
Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr Leu Gly
                565                 570                 575
Leu Asp Val Pro Val His His His His His His
            580                 585
```

<210> 9
<211> 583
<212> PRT
<213> Homo sapiens

<400> 9

```
Met Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg Arg Leu
                  5                  10                  15
Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Ala Met
             20                  25                  30
Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu Arg Lys
         35                  40                  45
Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Ile
     50                  55                  60
Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile Lys Val
 65                  70                  75                  80
Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu
                 85                  90                  95
Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg Leu Leu
            100                 105                 110
Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu Met Pro
        115                 120                 125
Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg Leu Gly
    130                 135                 140
Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly Met Ser
145                 150                 155                 160
Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn
                165                 170                 175
Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly Leu
            180                 185                 190
Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp Gly Gly
        195                 200                 205
Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg Arg Arg
    210                 215                 220
Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu
225                 230                 235                 240
Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala Arg Glu
                245                 250                 255
Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile
            260                 265                 270
Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile Asp
        275                 280                 285
Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe Ser Arg
    290                 295                 300
Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu Asp Leu
305                 310                 315                 320
```

```
Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu Leu Glu
                325                 330                 335
Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu Val Pro
            340                 345                 350
Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly Gly Met
        355                 360                 365
Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly Gly Asp
    370                 375                 380
Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg Ser Pro
385                 390                 395                 400
Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly Asp Leu
                405                 410                 415
Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His Asp Pro
            420                 425                 430
Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu Pro Ser
        435                 440                 445
Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln Pro Glu
    450                 455                 460
Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro Arg Glu
465                 470                 475                 480
Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Pro
                485                 490                 495
Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala
            500                 505                 510
Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly
        515                 520                 525
Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp
    530                 535                 540
Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro Pro
545                 550                 555                 560
Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr Leu Gly
                565                 570                 575
Leu Asp Val Pro Val Leu Glu
            580
```

<210> 10
<211> 589
<212> PRT
<213> Homo sapiens

<400> 10

```
Met Gln His His His His His His His Lys Arg Arg Gln Gln Lys Ile
                  5                  10                  15
Arg Lys Tyr Thr Met Arg Arg Leu Leu Gln Glu Thr Glu Leu Val Glu
             20                  25                  30
Pro Leu Thr Pro Ser Gly Ala Met Pro Asn Gln Ala Gln Met Arg Ile
         35                  40                  45
Leu Lys Glu Thr Glu Leu Arg Lys Val Lys Val Leu Gly Ser Gly Ala
     50                  55                  60
Phe Gly Thr Val Tyr Lys Gly Ile Trp Ile Pro Asp Gly Glu Asn Val
 65                  70                  75                  80
Lys Ile Pro Val Ala Ile Lys Val Leu Arg Glu Asn Thr Ser Pro Lys
                 85                  90                  95
Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met Ala Gly Val Gly
            100                 105                 110
Ser Pro Tyr Val Ser Arg Leu Leu Gly Ile Cys Leu Thr Ser Thr Val
        115                 120                 125
Gln Leu Val Thr Gln Leu Met Pro Tyr Gly Cys Leu Leu Asp His Val
    130                 135                 140
Arg Glu Asn Arg Gly Arg Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys
145                 150                 155                 160
```

```
Met Gln Ile Ala Lys Gly Met Ser Tyr Leu Glu Asp Val Arg Leu Val
                165                 170                 175
His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Ser Pro Asn His
            180                 185                 190
Val Lys Ile Thr Asp Phe Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu
        195                 200                 205
Thr Glu Tyr His Ala Asp Gly Gly Lys Val Pro Ile Lys Trp Met Ala
    210                 215                 220
Leu Glu Ser Ile Leu Arg Arg Arg Phe Thr His Gln Ser Asp Val Trp
225                 230                 235                 240
Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ala Lys Pro
                245                 250                 255
Tyr Asp Gly Ile Pro Ala Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly
            260                 265                 270
Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile
        275                 280                 285
Met Val Lys Cys Trp Met Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg
    290                 295                 300
Glu Leu Val Ser Glu Phe Ser Arg Met Ala Arg Asp Pro Gln Arg Phe
305                 310                 315                 320
Val Val Ile Gln Asn Glu Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser
                325                 330                 335
Thr Phe Tyr Arg Ser Leu Leu Glu Asp Asp Asp Met Gly Asp Leu Val
            340                 345                 350
Asp Ala Glu Glu Tyr Leu Val Pro Gln Gln Gly Phe Phe Cys Pro Asp
        355                 360                 365
Pro Ala Pro Gly Ala Gly Gly Met Val His His Arg His Arg Ser Ser
    370                 375                 380
Ser Thr Arg Ser Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser
385                 390                 395                 400
Glu Glu Glu Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly
                405                 410                 415
Ser Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu
            420                 425                 430
Gln Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu
        435                 440                 445
Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro
    450                 455                 460
Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro Asp Val Arg
465                 470                 475                 480
Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro
                485                 490                 495
Ala Gly Ala Thr Leu Glu Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn
            500                 505                 510
Gly Val Val Lys Asp Val Phe Ala Phe Gly Gly Ala Val Glu Asn Pro
        515                 520                 525
Glu Tyr Leu Thr Pro Gln Gly Gly Ala Ala Pro Gln Pro His Pro Pro
    530                 535                 540
Pro Ala Phe Ser Pro Ala Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp
545                 550                 555                 560
Pro Pro Glu Arg Gly Ala Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr
                565                 570                 575
Ala Glu Asn Pro Glu Tyr Leu Gly Leu Asp Val Pro Val
            580                 585
```

<210> 11
<211> 600
<212> PRT
<213> Homo sapiens

<400> 11

```
Met Gly His His His His His His His His Ser Ser Gly Ala Leu Asp
                  5                  10                  15
Asp Asp Asp Lys Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met
             20                  25                  30
Arg Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser
         35                  40                  45
Gly Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu
     50                  55                  60
Leu Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr
 65                  70                  75                  80
Lys Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala
                 85                  90                  95
Ile Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile
            100                 105                 110
Leu Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser
        115                 120                 125
Arg Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln
    130                 135                 140
Leu Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly
145                 150                 155                 160
Arg Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys
                165                 170                 175
Gly Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala
            180                 185                 190
Ala Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp
        195                 200                 205
Phe Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala
    210                 215                 220
Asp Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu
225                 230                 235                 240
Arg Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr
                245                 250                 255
Val Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro
            260                 265                 270
Ala Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln
        275                 280                 285
Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp
    290                 295                 300
Met Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu
305                 310                 315                 320
Phe Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn
                325                 330                 335
Glu Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser
            340                 345                 350
Leu Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr
        355                 360                 365
Leu Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala
    370                 375                 380
Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly
385                 390                 395                 400
Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro
                405                 410                 415
Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp
            420                 425                 430
Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr
        435                 440                 445
His Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro
    450                 455                 460
Leu Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro
465                 470                 475                 480
Gln Pro Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser
                485                 490                 495
```

```
Pro Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu
            500                 505                 510
Glu Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp
        515                 520                 525
Val Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro
    530                 535                 540
Gln Gly Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro
545                 550                 555                 560
Ala Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly
                565                 570                 575
Ala Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu
            580                 585                 590
Tyr Leu Gly Leu Asp Val Pro Val
        595                 600
```

<210> 12
<211> 957
<212> DNA
<213> Homo sapiens

<400> 12
atgggccccc agctccttgg ctatgtggtc ctttgccttc taggagcagg ccccctggaa 60
gcccaagtga cccagaaccc aagatacctc atcacagtga ctggaaagaa gttaacagtg 120
acttgttctc agaatatgaa ccatgagtat atgtcctggt atcgacaaga cccagggctg 180
ggcttaaggc agatctacta ttcaatgaat gttgaggtga ctgataaggg agatgttcct 240
gaagggtaca aagtctctcg aaaagagaag aggaatttcc ccctgatcct ggagtcgccc 300
agccccaacc agacctctct gtacttctgt gccagcagtt tagattgggg cggactagcg 360
ggagggttgg gcacagatac gcagtatttt ggcccaggca cccggctgac agtgctcgag 420
gacctgaaaa acgtgttccc acccgaggtc gctgtgtttg agccatcaga agcagagatc 480
tcccacaccc aaaaggccac actggtatgc ctggccacag gcttctaccc cgaccacgtg 540
gagctgagct ggtgggtgaa tgggaaggag gtgcacaagt ggggtcagca cagacccgca 600
gcccctcaag gagcaagccc gccctcaatg actccagata ctgctgagca gccgcctgag 660
ggtctcggcc acttctggca gaacccccgc aaccacttcc gctgtcaagt ccagttctac 720
gggctctcgg agaatgacga gtggacccag gatagggcca aacctgtcac ccagatcgtc 780
agcgccgagg cctggggtag agcagactgt ggcttcacct ccgagtctta ccagcaaggg 840
gtcctgtctg ccaccatcct ctatgagatc ttgctaggga aggccacctt gtatgccgtg 900
ctggtcagtg ccctcgtgct gatggccatg gtcaagagaa aggattccag aggctag   957

<210> 13
<211> 686
<212> DNA
<213> Homo sapiens

<400> 13
atggcctctg cacccatctc gatgcttgcg atgctcttca cattgagtgg gctgagagct 60
cagtcagtgg ctcagccgga agatcaggtc aacgttgctg aagggaatcc tctgactgtg 120
aaatgcacct attcagtctc tggaaaccct tatcttttt ggtatgttca ataccccaac 180
cgaggcctcc agttccttct gaaatacatc acaggggata acctggttaa aggcagctat 240
ggctttgaag ctgaatttaa caagagccaa acctccttcc acctgaagaa accatctgcc 300
cttgtgagcg actccgcttt gtacttctgt gctgtgagac cgaattcagg atacagcacc 360
ctcacctttg ggaaggggac tatgcttcta gtctctccag atatccagaa ccctgaccct 420
gccgtgtacc agctgagaga ctctaaatcc agtgacaagt ctgtctgcct attcaccgat 480
tttgattctc aaacaaatgt gtcacaaagt aaggattctg atgtgtatat cacagacaaa 540
actgtgctag acatgaggtc tatggacttc aagagcaaca gtgctgtggc ctggagcaac 600
aaatctgact ttgcatgtgc aaacgccttc aacaacagca ttattccaga agacaccttc 660
ttccccagcc cagaaagttc ctgtga   686

<210> 14
<211> 318
<212> PRT
<213> Homo sapiens

<400> 14
Met Gly Pro Gln Leu Leu Gly Tyr Val Val Leu Cys Leu Leu Gly Ala
1 5 10 15
Gly Pro Leu Glu Ala Gln Val Thr Gln Asn Pro Arg Tyr Leu Ile Thr
20 25 30
Val Thr Gly Lys Lys Leu Thr Val Thr Cys Ser Gln Asn Met Asn His
35 40 45
Glu Tyr Met Ser Trp Tyr Arg Gln Asp Pro Gly Leu Gly Leu Arg Gln
50 55 60
Ile Tyr Tyr Ser Met Asn Val Glu Val Thr Asp Lys Gly Asp Val Pro
65 70 75 80
Glu Gly Tyr Lys Val Ser Arg Lys Glu Lys Arg Asn Phe Pro Leu Ile
85 90 95
Leu Glu Ser Pro Ser Pro Asn Gln Thr Ser Leu Tyr Phe Cys Ala Ser
100 105 110
Ser Leu Asp Trp Gly Gly Leu Ala Gly Gly Leu Gly Thr Asp Thr Gln
115 120 125
Tyr Phe Gly Pro Gly Thr Arg Leu Thr Val Leu Glu Asp Leu Lys Asn
130 135 140
Val Phe Pro Pro Glu Val Ala Val Phe Glu Pro Ser Glu Ala Glu Ile
145 150 155 160
Ser His Thr Gln Lys Ala Thr Leu Val Cys Leu Ala Thr Gly Phe Tyr
165 170 175
Pro Asp His Val Glu Leu Ser Trp Trp Val Asn Gly Lys Glu Val His
180 185 190
Lys Trp Gly Gln His Arg Pro Ala Ala Pro Gln Gly Ala Ser Pro Pro
195 200 205
Ser Met Thr Pro Asp Thr Ala Glu Gln Pro Pro Glu Gly Leu Gly His
210 215 220
Phe Trp Gln Asn Pro Arg Asn His Phe Arg Cys Gln Val Gln Phe Tyr
225 230 235 240
Gly Leu Ser Glu Asn Asp Glu Trp Thr Gln Asp Arg Ala Lys Pro Val
245 250 255
Thr Gln Ile Val Ser Ala Glu Ala Trp Gly Arg Ala Asp Cys Gly Phe
260 265 270
Thr Ser Glu Ser Tyr Gln Gln Gly Val Leu Ser Ala Thr Ile Leu Tyr
275 280 285
Glu Ile Leu Leu Gly Lys Ala Thr Leu Tyr Ala Val Leu Val Ser Ala
290 295 300
Leu Val Leu Met Ala Met Val Lys Arg Lys Asp Ser Arg Gly
305 310 315

<210> 15
<211> 228
<212> PRT
<213> Homo sapiens

<400> 15
Met Ala Ser Ala Pro Ile Ser Met Leu Ala Met Leu Phe Thr Leu Ser
1 5 10 15
Gly Leu Arg Ala Gln Ser Val Ala Gln Pro Glu Asp Gln Val Asn Val
20 25 30
Ala Glu Gly Asn Pro Leu Thr Val Lys Cys Thr Tyr Ser Val Ser Gly
35 40 45
Asn Pro Tyr Leu Phe Trp Tyr Val Gln Tyr Pro Asn Arg Gly Leu Gln
50 55 60
Phe Leu Leu Lys Tyr Ile Thr Gly Asp Asn Leu Val Lys Gly Ser Tyr
65 70 75 80
Gly Phe Glu Ala Glu Phe Asn Lys Ser Gln Thr Ser Phe His Leu Lys
85 90 95

```
Lys Pro Ser Ala Leu Val Ser Asp Ser Ala Leu Tyr Phe Cys Ala Val
            100                 105                 110
Arg Pro Asn Ser Gly Tyr Ser Thr Leu Thr Phe Gly Lys Gly Thr Met
        115                 120                 125
Leu Leu Val Ser Pro Asp Ile Gln Asn Pro Asp Pro Ala Val Tyr Gln
    130                 135                 140
Leu Arg Asp Ser Lys Ser Ser Asp Lys Ser Val Cys Leu Phe Thr Asp
145                 150                 155                 160
Phe Asp Ser Gln Thr Asn Val Ser Gln Ser Lys Asp Ser Asp Val Tyr
                165                 170                 175
Ile Thr Asp Lys Thr Val Leu Asp Met Arg Ser Met Asp Phe Lys Ser
            180                 185                 190
Asn Ser Ala Val Ala Trp Ser Asn Lys Ser Asp Phe Ala Cys Ala Asn
        195                 200                 205
Ala Phe Asn Asn Ser Ile Ile Pro Glu Asp Thr Phe Phe Pro Ser Pro
    210                 215                 220
Glu Ser Ser Cys
225
```

<210> 16
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> primer PDM-44

<400> 16
atctctggcg cgctggatga cgatgacaag aaacgacggc agcagaag 48

<210> 17
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> primer PDM-45

<400> 17
cagggcgcgc cactcgagtc attacactgg cacgtccaga cccag 45

<210> 18
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> primer PDM-591

<400> 18
cacaaacgac ggcagcagaa gatccggaag 30

<210> 19
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> primer PDM-592

<400> 19

gcgccactcg agtcattaca ctggcacgtc 30

<210> 20
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> primer PDM-72

<400> 20
cgacttcata tgaaacgacg gcagcagaag atc 33

<210> 21
<211> 77
<212> DNA
<213> Artificial Sequence

<220>
<223> primer PDM-61

<400> 21
ccacgtctag agaaggcgcg ccatctggat cattaatgat gatgatgatg atgcactggc 60
acgtccagac ccaggta 77

<210> 22
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> primer TCR Valpha-16 5'

<400> 22
ggatccgccg ccaccatggc ctctgcaccc atctcga 37

<210> 23
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> primer TCR alpha 3'

<400> 23
gtcgactcag ctggaccaca gccgcag 27

<210> 24
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> primer TCR Vbeta-14. 5'

<400> 24
ggatccgccg ccaccatggg cccccagctc cttggcta 38

<210> 25
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> primer TCR beta 3'

<400> 25
gtcgactcag aaatcctttc tcttgac 27

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 February 2002 (21.02.2002)

PCT

(10) International Publication Number
WO 02/014503 A3

(51) International Patent Classification[7]: C12N 15/00, C07K 14/47

(21) International Application Number: PCT/US01/41733

(22) International Filing Date: 14 August 2001 (14.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/225,152	14 August 2000 (14.08.2000)	US
60/236,428	28 September 2000 (28.09.2000)	US
60/270,520	21 February 2001 (21.02.2001)	US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and

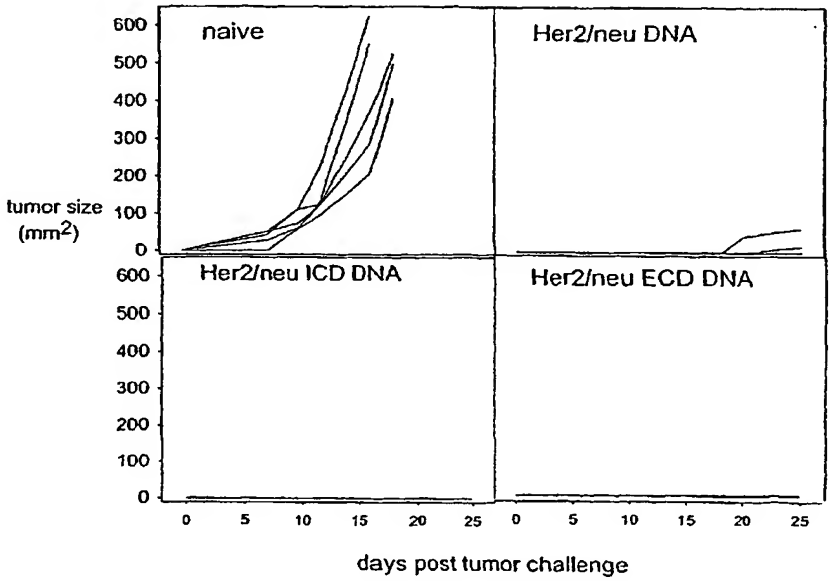
(75) Inventors/Applicants *(for US only)*: **HAND-ZIMMERMANN, Susan** [US/US]; 2014 179th Court N.E., Redmond, WA 98052 (US). **CHEEVER, Martin, A.** [US/US]; 6210 S.E. 22nd Avenue, Mercer Island, WA 98040 (US). **FOY, Teresa, M.** [US/US]; 2104 S. 277th Place, Federal Way, WA 98003 (US). **LODES, Michael, J.** [US/US]; 9223 36th Avenue S.W., Seattle, WA 98126 (US). **KALOS, Michael, D.** [US/US]; 8116 Dayton Avenue N., Seattle, WA 98103 (US). **MCNEILL, Patricia, D.** [US/US]; 1333 South 290th Place, Federal Way, WA 98003 (US). **VEDVICK, Thomas, S.** [US/US]; 124 S. 300th Place, Federal Way, WA 98003 (US).

(74) Agents: **CHRISTIANSEN, William, T.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF HER-2/NEU-ASSOCIATED MALIGNANCIES



(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly Her-2/neu-associated cancers, are disclosed. Illustrative compositions comprise one or more Her-2/neu polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of Her-2/neu-associated malignancies.

WO 02/014503 A3

SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
18 September 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/41733

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/00 C07K14/47

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 45954 A (EPIMMUNE INC) 16 September 1999 (1999-09-16) see claims and Table 3	1-10
X	WO 00 44899 A (GHEYSEN DIRK ;SMITHKLINE BEECHAM (GB); CORIXA CORP (US); CHEEVER M) 3 August 2000 (2000-08-03) see the whole document, especially Claim 27 and Figure 11	1-10
X	WO 93 14781 A (UNIV CALIFORNIA) 5 August 1993 (1993-08-05) see Claim 22 and page 70	1-8
Y	US 5 801 005 A (DISIS MARY L ET AL) 1 September 1998 (1998-09-01) column 11, line 37	1-10

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 November 2002	05. 03. 03

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Grosskopf, R

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/41733

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 99 57981 A (SLOAN KETTERING INST CANCER ;ZELENETZ ANDREW D (US); ROBERTS WENDY) 18 November 1999 (1999-11-18) see the whole document	1-10
A	CHEN YING ET AL: "DNA vaccines encoding full-length or truncated Neu induce protective immunity against Neu-expressing mammary tumors" CANCER RESEARCH, AMERICAN ASSOCIATION FOR CANCER RESEARCH, BALTIMORE, MD, US, vol. 58, no. 9, 1 May 1998 (1998-05-01), pages 1965-1971, XP002149613 ISSN: 0008-5472	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/41733

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 6 to 10 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-10

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-10

Isolated polynucleotide encoding the peptide having the amino acid sequence of SEQ ID NO: 3, the corresponding peptide, pharmaceutical compositions comprising said entities and methods for elicting an immune response.

2. Claim : 11

An isolated polynucleotide composition comprising the TCR-alpha sequence set forth in SEQ ID NO: 13

3. Claim : 12

An isolated polynucleotide composition comprising the TCR-beta sequence set forth in SEQ ID NO: 12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/41733

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9945954 A	16-09-1999	AU 6465598 A	27-09-1999
		CA 2323632 A	16-09-1999
		EP 1064022 A	03-01-2001
		JP 2002507397 T	12-03-2002
WO 0044899 A	03-08-2000	AU 2742600 A	18-08-2000
		BR 0007840 A	22-01-2002
		CA 2361009 A	03-08-2000
		CN 1345374 T	17-04-2002
		CZ 20012587 A	15-05-2002
		EP 1147190 A	24-10-2001
		HU 0105303 A	29-05-2002
		JP 2002535004 T	22-10-2002
		NO 20013701 A	28-09-2001
		TR 200102191 T	21-12-2001
		US 2002177567 A	28-11-2002
WO 9314781 A	05-08-1993	NONE	
US 5801005 A	01-09-1998	AU 708237 B	29-07-1999
		AU 5522296 A	16-10-1996
		BR 9607889 A	10-08-1999
		CA 2216601 A	03-10-1996
		CN 1183117 A	27-05-1998
		CZ 9703096 A	14-01-1998
		EP 0817846 A	14-01-1998
		HU 9801826 A	30-11-1998
		JP 11502710 T	09-03-1999
		NO 974502 A	27-11-1997
		NZ 306616 A	30-03-2001
		WO 9630514 A	03-10-1996
		US 2002055614 A	09-05-2002
		US 6075122 A	13-06-2000
		US 5726023 A	10-03-1998
		US 5876712 A	02-03-1999
		US 5846538 A	08-12-1998
		US 5869445 A	09-02-1999
WO 9957981 A	18-11-1999	EP 1075184 A	14-02-2001
		JP 2002514573 T	21-05-2002